
Stochastic models in financial risk management

Von der Fakultät für MINT – Mathematik, Informatik, Physik,
Elektro- und Informationstechnik

der Brandenburgischen Technischen Universität Cottbus–Senftenberg

genehmigte

Dissertation

zur Erlangung des akademischen Grades eines

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt von

Imke Redeker, geborene Höfers

geboren am 11.08.1989 in Wilhelmshaven

Vorsitzende/r: Prof. Dr. Sabine Pickenhain

Gutachter/in: Prof. Dr. Ralf Wunderlich

Gutachter/in: Prof. Dr. Michaela Szölgényi

Tag der mündlichen Prüfung: 15.03.2019

Kurzfassung

In dieser Arbeit werden stochastische Modelle zur Betrachtung von Markt- und Kreditrisiken untersucht. Der erste Teil befasst sich mit dem Marktrisiko und untersucht optimale Investitions- und Konsumstrategien für einen Investor am Finanzmarkt, der angehalten ist, seine (Anlage-) Strategie so zu wählen, dass das Risiko für den Verlust eines Teils seines investierten Kapitals beschränkt bleibt. Die Beschränkung des Verlustrisikos erfolgt mittels dynamischer Risikomaße, die zu jedem Zeitpunkt das Risiko unter Berücksichtigung der bis zu diesem Zeitpunkt vorliegenden Information quantitativ erfassen. Dabei werden die dynamischen Risikomaße Value at Risk, Tail Conditional Expectation und Expected Loss verwendet. Die entstehenden Portfolio-Probleme unter dynamischen Risikoebenenbedingungen werden sowohl in stetiger als auch in diskreter Zeit betrachtet. Zur Lösung der Portfolio-Probleme werden diese als stochastische Kontrollprobleme formuliert und mit Methoden des Dynamic Programming gelöst. Der Verlust der Portfolio-Performance infolge der Beschränkung des Risikos sowie der Verlust durch die Beschränkung auf zeitdiskrete Investmentstrategien werden untersucht und eine Vielzahl von numerischen Experimenten wird präsentiert. Diese zeigen, dass der Verlust der Portfolio-Performance nicht erheblich ist, während das Risiko deutlich beschränkt wird. Darüber hinaus ist der Verlust, der durch die Beschränkung auf zeitdiskrete Investmentstrategien entsteht, typischerweise größer als der Verlust der Portfolio-Performance durch eine Risikobeschränkung.

Der zweite Teil dieser Arbeit befasst sich mit dem Kreditrisiko. Zur Modellierung des Risikos, dass ein Kreditnehmer seinen vertragskonformen Zahlungsverpflichtungen nicht oder nur teilweise nachkommt, wird ein Modell untersucht, das die beiden klassischen Ansätze zur Modellierung des Kreditrisikos kombiniert. In diesem Modell erfolgt der Kreditausfall eines Unternehmens bei Unterschreitung eines Schwellenwertes durch das Unternehmensvermögen. Der Schwellenwert wird dabei als zeitabhängige Zufallsvariable modelliert, die es erlaubt, Veränderungen in der Wirtschaft oder die Ernennung einer neuen Unternehmensleitung abzubilden. Es werden in Abhängigkeit von der verfügbaren Information über das Unternehmensvermögen explizite Formeln für die Ausfallwahrscheinlichkeit hergeleitet. Der Einfluss asymmetrischer Informationen auf die Ausfallwahrscheinlichkeit und die (Ausfall-) Risikoprämie wird untersucht. Numerische Ergebnisse werden präsentiert, welche zeigen, dass die Informationen über den Unternehmenswert einen erheblichen Einfluss auf die Schätzung der Ausfallwahrscheinlichkeit und der damit einhergehenden Risikoprämie haben.

Abstract

This thesis is concerned with stochastic models to manage financial risks. The first part deals with market risk and considers an investor facing a classical portfolio problem of optimal investment in log-Brownian stocks and a fixed-interest bond, but constrained to choose portfolio and consumption strategies which reduce the corresponding shortfall risk. Risk limits are formulated in terms of Value at Risk, Tail Conditional Expectation and Expected Loss and are dynamically imposed on the strategy as a risk constraint. The resulting stochastic optimal control problem is tackled using the dynamic programming approach. For both continuous-time and discrete-time financial markets the loss in expected utility of intermediate consumption and terminal wealth caused by imposing a dynamic risk constraint is investigated. The presented numerical results indicate that the loss of portfolio performance is not too large while the risk is notably reduced. Furthermore, the loss resulting from infrequent trading due to time discretization effects is typically bigger than the loss of portfolio performance resulting from imposing a risk constraint.

The second part deals with credit risk and sets up a first-passage model of corporate default risk. The default event is specified in terms of the evolution of the total value of the firm's asset and the default barrier. Short-term default risk is incorporated by modeling the default barrier at which the firm is liquidated as a random variable which is time-dependent and allowed to switch. This setup combines the two classical modeling approaches and enables to model changes in the economy or the appointment of a new firm management. Different information levels on the firm's assets are distinguished and explicit formulas for the conditional default probability given the accessible information are derived. The impact of asymmetric information on the default probability and credit yield spread is investigated. Numerical results are presented indicating that the information on the firm value has a considerable impact on the estimate of the conditional survival probability and the associated credit yield spread.

Contents

1	Introduction	1
2	Portfolio optimization under dynamic risk constraints	7
2.1	Continuous-time optimization	8
2.1.1	Model	8
2.1.2	Dynamic risk constraints	10
2.1.3	Optimization under risk constraints	13
2.2	Discrete-time optimization	15
2.2.1	Model	15
2.2.2	Dynamic risk constraints	17
2.2.3	Optimization under risk constraints	18
2.3	Numerical examples for power utility	22
2.3.1	Continuous-time optimization problem	22
2.3.2	Discrete-time optimization problem	28
2.3.3	Financial market with a bond and a single stock	28
2.3.4	Financial market with a bond and two stocks	36
2.4	Proofs	53
2.4.1	Proof of Lemma 2.1.3	53
2.4.2	Proof of Lemma 2.2.1	54
2.5	Results from dynamic programming	58
2.5.1	Continuous time	58
2.5.2	Discrete time	64
3	Credit risk with asymmetric information and a switching barrier	69
3.1	Credit risk models	69
3.1.1	Basic definitions and notation	70
3.2	Model for the default event	73
3.2.1	Default barrier	73
3.2.2	Information structure	74
3.3	Conditional survival probability	76
3.3.1	Preliminaries	76
3.3.2	S-investor's information case	77

3.3.3	D-investor's information case	81
3.4	Numerical examples	83
3.4.1	Case of a constant default threshold	83
3.4.2	Case of a switching default threshold	89
3.5	Proofs	96
3.5.1	Proof of Lemma 3.3.2	96
3.5.2	Proof of Equation (3.6)	96
3.5.3	Proof of Lemma 3.3.6	99
3.5.4	Proof of Theorem 3.3.7	100
4	Conclusions	106
	Bibliography	108
	List of Figures	114
	Notation	116

Chapter 1

Introduction

Financial risk management is a process of measuring and mitigating financial risks and has become a core competence of financial institutions. The challenge is to effectively manage the risk on the one hand and ensure adequate returns on the other. Due to globalization financial business is becoming more and more complicated and there is a great need to expand modeling techniques in the financial industry to manage the resulting risks. The first step in risk management is to identify the sources of potential losses. The broad categories of financial risk are market risk, credit risk, and operational risk (see MCNEIL, FREY & EMBRECHTS [58]). Market risk is the risk of a change in the value of a financial position. Credit risk corresponds to the risk to each party of a contract that the other will fail to comply with its obligation to service debt. Operational risk refers to the risk of losses resulting from inadequate or failed internal processes, people, and systems, or from external events. This thesis is specifically concerned with market risk and credit risk.

When trading assets in a financial market, one has to consider the risk of fluctuations of asset prices due to market volatility, i.e., market risk. An investor who is endowed with a certain initial wealth may consume parts of the wealth and continuously invest in different financial assets. His preferences are represented by utility functions. Given a fixed investment horizon, the investor's objective is to find an optimal consumption-investment strategy as to maximize the expected utility of wealth at the terminal trading time and of intermediate consumption. This type of continuous-time portfolio optimization problem was first studied by MERTON [59, 60] in the late sixties and early seventies by applying standard methods and results from stochastic control theory.

In [59] the author considers a financial market consisting of one risk-free and several risky assets. The prices of the risky assets are modeled by a stochastic process, namely a geometric Brownian motion, with dynamics evolving according to a linear stochastic differential equation. This model assumption leads to prices of risky assets that are lognormally distributed random variables and to price processes with

continuous paths. The investor's utility function is modeled by power utility. Using methods of stochastic optimal control theory by applying the dynamic programming principle and solving the resulting Hamilton-Jacobi-Bellman (HJB) equation, a nonlinear partial differential equation, Merton derives an explicit formula for the consumption-investment strategy. Thus, the Merton approach must be regarded as the starting point of continuous-time portfolio theory and much recent work deals with refined models and methods in the field of portfolio optimization. An introduction to stochastic optimal control theory can be found in FLEMING & SONER [33], ØKSENDAL & SULEM [65], and PHAM [67].

The financial crisis has drawn more attention to risk management by traders and regulators in the past years. Since then, various risk measures like Value at Risk (VaR) or some tail-expectation-based risk measures like Tail Conditional Expectation (TCE) or Expected Loss (EL) have been increasingly applied to quantify and control the risk of a portfolio. VaR is widely used in risk management since it can be well understood and communicated due to its simplicity. Usually risk constraints having been considered are static constraints, i.e., they have to hold at the terminal trading time only (see e.g., BASAK & SHAPIRO [5], BOYLE & TIAN [21]). In the seminal paper [5] the authors compute the VaR of the terminal wealth to control the risk exposure. Their findings indicate that VaR limits, when applied in a static manner, may actually increase risk. This encouraged researchers to consider a risk measure that is based on the risk neutral expectation of loss - the Limited Expected Loss. The work of [5] is extended by GABIH ET AL. [35] to cover the case of bounded Expected Loss.

Motivated by the Basel Committee proposals, it is a common practice in the financial industry to compute and re-evaluate risk constraints frequently using a time window (e.g., some days or weeks) over which the trading strategies are assumed to be held constant (see JORION [49]). Therefore CUOCO ET AL. [24] apply a risk constraint dynamically to control the risk exposure. The authors find that VaR and TCE constraints, when applied in a dynamic fashion, reduce the investment in the risky asset. The work of [24] is extended by SHI & WERKER [76] to cover the case of stochastic interest rates. The authors analyze the impacts of imposing VaR and expected shortfall constraints and show that these constraints, when dynamically imposed, lead to similar optimal portfolios and wealth distributions. PUTSCHÖGL & SASS [71] study the maximization of expected utility of terminal wealth under dynamic risk constraints in a complete market model with partial information on the drift by using the martingale method. The problem of intertemporal consumption subject to a dynamic VaR constraint is examined by YIU [82]. PIRVU [68] and LEIPPOLD ET AL. [55] investigate an optimal consumption-investment problem with a dynamic VaR constraint imposed on the strategy. AKUME [1] and AKUME ET AL. [2] consider a similar problem with a dynamic TCE constraint instead of a VaR. Their results indicate that imposing a dynamic risk constraint is a suitable method to reduce the risk of portfolios. For recent work on portfolio optimization under dynamic risk con-

straints we refer to MORENO-BROMBERG ET AL. [69]. The authors study an optimal investment problem for a general class of risk measures and the optimal trading strategy is characterized by a quadratic backward stochastic differential equation. For the purpose of dynamic risk measurement, the usual assumption for a tractable way of calculating risk measures is that the investment and consumption strategies are kept unchanged over a given time horizon. In the classical consumption-investment problem, this assumption is typically not fulfilled. An investor who follows the optimal strategy is continuously adjusting the asset holdings and the consumption rate. Thus, the risk of a portfolio is only approximately calculated. Since continuous-time trading is impossible in practice and investors want to avoid transaction costs due to excessive trading a more realistic scenario is that investors only change their asset holdings and consumption rate at discrete points in time. In ROGERS [74] and BÄUERLE ET AL. [9] such discrete-time investors are called relaxed investors. It is shown that the losses due to discretization are surprisingly small.

In all of the above mentioned work in the field of portfolio optimization, only market risk is considered. However, another type of risk the financial industry may have to deal with is credit risk. When a counterparty is involved in a financial transaction, one has to consider the risk that the counterparty may not fulfill its contractually agreed payment obligations, i.e., credit risk. A credit risky (or equivalently default risky) asset is a security that has a nonzero probability of defaulting on its contracted payments. A typical example of a credit risky asset is a corporate bond. A corporate bond promises its holder a fixed stream of payments but may default on its promise. The role of a typical credit risk model is to describe the default of a default risky asset. The recent financial crisis has impressively demonstrated the need for effective credit risk management. Since then the evaluation of credit risk has been receiving increasing attention. For credit risk analysis the modeling of a default event and the forecasting of a default probability is crucial both from a theoretical and an empirical point of view. There are two classical types of modeling approaches: the structural one and the reduced-form one. The structural approach is considered by BLACK & SCHOLES [15], MERTON [61] and BLACK & COX [14], among others. It provides a relationship between default risk and capital structure by using the evolution of the firm's assets value to determine the time of default, i.e., the default event of a bond is triggered when the assets of the firm who issued the bond fall below some threshold. The important feature of the structural model is that it implicitly assumes that the modeler has complete knowledge about the dynamics of the firm's assets and the situation that will trigger the default event (i.e., the firm's liabilities). Despite the convincing economic interpretation in terms of the firm's assets and liabilities there are shortcomings when the firm's assets are modeled by a continuous-time asset value process. One is that credit yield spreads go to zero as maturity goes to zero regardless of the riskiness of the firm. This results from the investors' knowledge about the firm's true distance to default. Such credit spreads are uncommon in practice.

Another disadvantage of the structural approach is that forecast bond prices continuously converge to their recovery value (the payment which is received if default occurs before maturity) which contradicts the price jump at default in empirical studies. These issues do not occur in the second approach, the reduced-form approach, which is considered by JARROW & TURNBULL [46], ARTZNER & DELBAEN [3], and DUFFIE & SINGLETON [28], among others. It treats the dynamics of default as an exogenous event. This implies knowledge of a less detailed information set compared to the structural approach and credit spreads become in general more realistic and are easier to quantify. Another advantage is that the reduced-form approach has proven to be very useful for the valuation of credit-sensitive securities. However, the approach is lacking of economic insights as it does not connect credit risk to underlying structural variables. To gain both the economic appeal of the structural approach and the empirical plausibility and the tractability of the reduced-form approach, structural models can be transformed into reduced-form models by changing its information set to a less refined one. One way is to model the default barrier as a random variable which is unobservable by bond investors (see LANDO [54], GIESECKE & GOLDBERG [36], HILLAIRET & JIAO [41]). Another way is to assume that the firm's assets are only partially observable by investors (see DUFFIE & LANDO [27], JEANBLANC & VALCHEV [47], LAKNER & LIANG [53]).

The first part of this thesis focuses on *market risk*. We consider an investor faced with the classical consumption-investment problem, but constrained to choose investment and consumption strategies which reduce the corresponding shortfall risk. Risk measures are approximately calculated for short time horizons and imposed on the strategy as a risk constraint. We apply dynamic programming techniques and combine the resulting Hamilton-Jacobi-Bellman equation with the method of Lagrange multipliers to derive optimal strategies under this constraint. An approximate solution to the constrained portfolio problem is obtained by using a policy improvement algorithm. This continuous-time investor is compared to a discrete-time investor who chooses only to change portfolio and consumption choices at times which are multiples of $\Delta > 0$. Within the framework of a discrete-time investor we consider a discretized standard Black-Scholes market with one risk-free and several risky assets. The discrete-time investor is, like the continuous-time investor, constrained to limit the risk exposure. The investor's aim is to maximize the expected utility of intermediate consumption and of wealth at the terminal trading time. The portfolio risk is re-evaluated dynamically making use of available conditioning information. Risk limits are formulated in terms of VaR, TCE, and EL. The optimization problem is solved by using the theory of Markov Decision Problems and numerical results are presented which illustrate the impact on the portfolio performance resulting from imposing a dynamic risk constraint as well as from infrequent trading in a discrete-time setting.

The second part of this thesis deals with *credit risk* and generalizes previous works

to a dynamic setting. The default event is specified in terms of the evolution of the total value of the firm's assets and the default threshold. Short-term default risk is incorporated by modeling the default threshold at which the firm is liquidated as a random variable whose value is chosen by the management of the firm and adjusted dynamically. In the literature, this generalization of the default model to a dynamic setting was proposed by BLANCHET-SCALLIET, HILLAIRET & JIAO [16]. The authors study the information accessible to the management of the firm and obtain explicit formulations for the survival probability given the information of the management by using a successive enlargement framework. Our approach is different and related to ordinary investors on the market who do not have access to the value of the threshold and only anticipate the distribution of the threshold. We distinguish different information levels on the firm's assets and derive explicit formulas for the conditional default probability given these information levels. Numerical results are presented which show the impact of the information level on the estimation of the default probability and the associated credit yield spread.

The first objective of the thesis refers to market risk and is to analyze how an investor who is endowed with a certain initial wealth has to consume and invest this wealth optimally in a financial market when he is subjected to fulfill a given risk constraint in a dynamic fashion. For both a continuous-time and a discrete-time financial market, different dynamic risk constraints are considered and the corresponding optimal consumption and investment strategy is investigated. The second objective refers to credit risk and is to analyze the impact of information accessibility of an investor on the estimation of the default probability within a generalized structural model for credit risk. The main contributions with regard to market risk are:

- We incorporate dynamic risk measures in a discrete-time financial market to overcome the disadvantage that the risk of a portfolio is only approximately calculated in a continuous-time setting.
- We implement numerical methods (meshless method, sequential quadratic programming, numerical integration) to solve the continuous-time and discrete-time portfolio optimization problem under dynamic risk constraints.
- We provide extensive numerical experiments to explore the loss of portfolio performance resulting from imposing a risk constraint in comparison to the loss resulting from infrequent trading. Furthermore, in a financial market with more than one stock the impact of the stocks' correlation coefficient on the optimal investment strategy is investigated.

The main contributions with regard to credit risk are:

- We extend the traditional structural model for credit risk to a dynamic setting by relaxing the assumption of a constant default threshold to the case of a piecewise constant threshold which is unknown to the investor. Explicit formulas for

the default probabilities given different information levels on the firm's assets are derived.

- We provide numerical examples to compare the survival probability estimated by an investor who only observes the firm value at discrete dates to the survival probability estimated by an investor who continuously observes the firm value and investigate the impact of discrete-time observations.

The thesis is organized in two main chapters. Chapter 2 deals with incorporating dynamic risk constraints in continuous- and discrete-time financial markets to quantify and control the risk of fluctuations of asset prices due to market volatility. This chapter extends the master's thesis of REDEKER [72] by considering discrete-time financial markets and providing numerical examples for financial markets with more than one stock. Parts of this chapter have already appeared in REDEKER & WUNDERLICH [73]. Chapter 3 is devoted to credit risky assets and concerned with modeling the default event and forecasting the default probability of a credit risky asset.

In Section 2.1 we start with considering a continuous-time investor facing a classical consumption-investment problem, but constrained to choose investment and consumption strategies that reduce a corresponding shortfall risk. The investor's aim is to maximize the expected utility of intermediate consumption and wealth at a terminal trading time. After considering the continuous-time case of portfolio optimization under dynamic risk constraints, Section 2.2 is devoted to the discrete-time case. In Section 2.3 numerical results are presented for both cases and the cost of imposing a risk constraint is analyzed as well as the loss due to time-discretization. In Section 2.4 we provide proofs omitted from the main text and Section 2.5 presents some results from dynamic programming in continuous time and discrete time.

In Chapter 3 we proceed as follows. Section 3.1 introduces the structural and reduced-form modeling approaches for credit risk. Section 3.2 sets up a model for credit risk based on a structural model but with an unobservable default barrier that is allowed to switch. In this set up different information structures are distinguished. Section 3.3 studies the impact of asymmetric information on the default probability. We derive explicit formulas for the conditional survival probabilities given the different information structures. Section 3.4 presents numerical results illustrating the impact of asymmetric information on the default probability. Section 3.5 contains proofs omitted from the main text.

The thesis concludes with Chapter 4 containing a summary of key findings and remarks on possible future research.

Chapter 2

Portfolio optimization under dynamic risk constraints

Market risk comprises the risk of a loss due to price movements in financial securities. In this chapter, we investigate dynamic risk measures and dynamically impose risk limits on the trading strategies to control market risk. The associated portfolio optimization problem of an investor is to find a strategy which optimally allocates his wealth between different financial securities and consumption on the one hand and satisfies the imposed risk constraint on the other hand. The investor's aim is to maximize the expected utility of the intermediate consumption and wealth at terminal trading time. This portfolio optimization problem is analyzed in continuous-time and discrete-time financial markets.

We start with considering the portfolio problem in continuous time. In Section 2.1.1 we describe the financial market model, define admissible consumption and investment strategies and derive the investor's wealth equation. Afterwards, in Section 2.1.2, we introduce dynamic risk constraints. The portfolio risk in terms of VaR, TCE, and EL is measured over short time intervals. The risk measure is dynamically re-evaluated using available conditioning information and imposed on the strategy as a risk constraint. Section 2.1.3 formulates the portfolio optimization problem under dynamic risk constraints. Dynamic programming techniques are applied and the resulting Hamilton-Jacobi-Bellman equation is combined with the method of Lagrange multipliers to derive optimal strategies under risk constraints.

Then, we consider the portfolio problem in discrete time. In Section 2.2.1 we consider a discretized version of the financial market previously introduced. An investor in this financial market only changes portfolio and consumption choices at discrete time points. However, the discrete-time investor is, like the continuous-time investor, constrained to limit the risk exposure. In Section 2.2.2 the dynamic risk constraints are described in discrete time. We subsequently study the discrete-time portfolio optimization problem under dynamic risk constraints in Section 2.2.3. The optimization problem is solved by using the theory of Markov Decision Problems leading to a

backward recursion algorithm.

Section 2.3 provides some numerical results of the continuous-time and discrete-time portfolio optimization problem for the case of power utility. The cost of the risk constraint in terms of the expected utility of intermediate consumption and terminal wealth is investigated. Furthermore, numerical experiments are performed to study the losses due to time-discretization.

Proofs omitted from the main text are presented in Section 2.4 and some results from dynamic programming in continuous time and discrete time are provided for convenience of the reader in Section 2.5.

2.1 Continuous-time optimization

2.1.1 Model

We consider a continuous-time stochastic financial market with finite trading horizon $0 < T < \infty$ and an investor who is endowed with an initial capital $x_0 > 0$. The investor's possible actions are to invest in the financial market and/or to consume (parts of) the wealth. The investment opportunities are represented by one risk-free and d risky securities. The price of the risk-free security at time t is denoted by S_t^0 and the d -dimensional price process of the risky securities is denoted by $\mathbf{S} = (\mathbf{S}_t)_{t \in [0, T]}$. Uncertainty is modeled by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ is the natural filtration generated by an m -dimensional Brownian motion $\mathbf{W} = (\mathbf{W}_t)_{t \in [0, T]}$, $m \geq d$, augmented by all the \mathbb{P} -null sets of Ω .

The risk-free security (the ‘‘bond’’) behaves like a bank account earning a continuously compounded interest rate $r \geq 0$, i.e., its price at time t is given by

$$S_t^0 = e^{rt}.$$

The remaining d securities (the ‘‘stocks’’) are risky and evolve according to the following stochastic differential equations

$$\begin{aligned} dS_t^i &= S_t^i \left(\mu^i dt + \sum_{j=1}^m \sigma^{ij} dW_t^j \right), \quad i = 1, \dots, d, \\ S_0^i &= s^i, \end{aligned}$$

where $\boldsymbol{\mu} = (\mu^i)_{i=1, \dots, d} \in \mathbb{R}^d$ is the mean rate of return and $\boldsymbol{\sigma} = (\sigma^{ij})_{i=1, \dots, d}^{j=1, \dots, m} \in \mathbb{R}^{d \times m}$ the matrix-valued volatility which is assumed to consist of linearly independent rows. Thus, we work on a standard Black-Scholes-Merton model. The bond and the stocks can be traded continuously and the investor is allowed to sell the stocks or the bond short. An investment strategy $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t \in [0, T]}$ is an \mathbb{F} -adapted, real-valued,

d -dimensional stochastic process satisfying

$$\mathbb{E} \left[\int_0^T \|\boldsymbol{\pi}_t\|^2 dt \right] < \infty,$$

where $\|\cdot\|$ denotes the standard Euclidean norm in \mathbb{R}^d . The d coordinates of $\boldsymbol{\pi}_t$ represent the proportions of the current wealth invested in each of the d stocks at time t .

A negative proportion π_t^i corresponds to selling stock i short. The proportion of wealth invested in the bond at time t is given by $1 - \sum_{i=1}^d \pi_t^i$. If this quantity is negative we sell the bond short, i.e., we are borrowing (at interest rate r). The investor is also allowed to withdraw funds for consumption. The consumption rate process (for brevity: consumption process) is denoted by $C = (C_t)_{t \in [0, T]}$. It is an adapted, $[0, \infty)$ -valued stochastic process satisfying

$$\mathbb{E} \left[\int_0^T C_t dt \right] < \infty.$$

Given an investment and consumption strategy the associated wealth process is well defined and satisfies the stochastic differential equation

$$\begin{aligned} dX_t &= X_t \left(1 - \sum_{i=1}^d \pi_t^i \right) r dt + X_t \boldsymbol{\pi}'_t \boldsymbol{\mu} dt - C_t dt + X_t \boldsymbol{\pi}'_t \boldsymbol{\sigma} d\mathbf{W}_t \\ &= X_t (r + \boldsymbol{\pi}'_t (\boldsymbol{\mu} - \mathbf{1}r) - c_t) dt + X_t \boldsymbol{\pi}'_t \boldsymbol{\sigma} d\mathbf{W}_t, \end{aligned} \quad (2.1)$$

where $\boldsymbol{\pi}'_t$ denotes the transpose of $\boldsymbol{\pi}_t$, $\mathbf{1} := (1, \dots, 1)'$ is the d -dimensional vector with unit components and $c_t = C_t/X_t$ is the consumption rate relative to the wealth. Together with the initial condition $X_0 = x_0$, the wealth equation (2.1) admits a unique strong solution given by

$$X_t = x_0 \exp \left\{ \int_0^t \left(r + \boldsymbol{\pi}'_s (\boldsymbol{\mu} - \mathbf{1}r) - c_s - \frac{1}{2} \|\boldsymbol{\pi}'_s \boldsymbol{\sigma}\|^2 \right) ds + \int_0^t \boldsymbol{\pi}'_s \boldsymbol{\sigma} d\mathbf{W}_s \right\}. \quad (2.2)$$

Note that (2.2) implies

$$X_{t+\Delta} = X_t \exp \left\{ \int_t^{t+\Delta} \left(r + \boldsymbol{\pi}'_s (\boldsymbol{\mu} - \mathbf{1}r) - c_s - \frac{1}{2} \|\boldsymbol{\pi}'_s \boldsymbol{\sigma}\|^2 \right) ds + \int_t^{t+\Delta} \boldsymbol{\pi}'_s \boldsymbol{\sigma} d\mathbf{W}_s \right\} \quad (2.3)$$

for any $\Delta > 0$. A pair $(\boldsymbol{\pi}_t, c_t)_{t \in [0, T]} = ((\pi_t^i)_{i=1, \dots, d}, c_t)_{t \in [0, T]}$ is called a portfolio-proportion process. We only consider those portfolio-proportion processes $\mathbf{u} = (\boldsymbol{\pi}, c)$ which achieve a positive wealth over the whole trading period $[0, T]$ and which are of Markov type, i.e, $\mathbf{u}_t = \tilde{u}(t, X_t)$ for all $t \in [0, T]$ and some measurable function

$\tilde{u}: [0, T] \times \mathbb{R}_+ \rightarrow \mathcal{K}$ where $\mathcal{K} := \mathbb{R}^d \times [0, \infty)$. Such strategies are called admissible and the set of admissible strategies is denoted by \mathcal{A}_0 , thus

$$\mathcal{A}_0 := \left\{ (\mathbf{u}_t)_{t \in [0, T]} \mid \mathbf{u} \text{ is } \mathbb{F}\text{-adapted, } \mathbf{u}_t = (\boldsymbol{\pi}_t, c_t) \in \mathcal{K}, \mathbf{u}_t = \tilde{u}(t, X_t), \right. \\ \left. X_t > 0 \text{ for all } t \in [0, T] \text{ and } \mathbb{E} \left[\int_0^T (\|\boldsymbol{\pi}_t\|^2 + c_t) dt \right] < \infty \right\}.$$

We write in the following $X^{\mathbf{u}}$ instead of X to emphasize that the wealth is controlled by the portfolio-proportion process $\mathbf{u} = (\boldsymbol{\pi}, c)$.

2.1.2 Dynamic risk constraints

In this section we introduce how the risk of a given portfolio-proportion process can be quantified. In ARTZNER ET AL. [4] risk is defined by the random future value of the portfolio wealth. In order to relate this definition of risk to the investor's loss we use the concept of benchmarks as in AKUME [1] and AKUME ET AL. [2].

Given the current time $t \in [0, T]$ a benchmark Y_t is prescribed and compared to the future portfolio value $X_{t+\Delta}^{\mathbf{u}}$ at time $t + \Delta$, where $\Delta > 0$ is the length of the risk measurement horizon $[t, t + \Delta]$. Then a shortfall is described by the random event $\{X_{t+\Delta}^{\mathbf{u}} < Y_t\}$ and $L_t := Y_t - X_{t+\Delta}^{\mathbf{u}}$ is the corresponding investor's loss. The benchmark Y_t is chosen as a function of time t and wealth $X_t^{\mathbf{u}}$ for all $t \in [0, T]$, i.e., $Y_t = \tilde{f}(t, X_t^{\mathbf{u}})$ for some measurable function $\tilde{f}: [0, T] \times \mathbb{R}_+ \rightarrow [0, \infty)$. Typical benchmarks are presented in the following example.

Example 2.1.1. The benchmark may be chosen as

- a constant, i.e., $Y_t \equiv y$ for some $y \geq 0$,
- a deterministic function, i.e., $Y_t = y(t)$,
- a fraction $p > 0$ of the current wealth, i.e., $Y_t = pX_t^{\mathbf{u}}$,
- the conditional expected wealth, i.e., $Y_t = \mathbb{E}[X_{t+\Delta}^{\mathbf{u}} | \mathcal{F}_t]$.

Next, we make precise how the risk of a given portfolio-proportion process is measured. Let

$$\mathcal{A} := \left\{ (\mathbf{u}_t)_{t \in [0, T]} \in \mathcal{A}_0 \mid \mathbb{E}[|L_t|] < \infty \text{ for all } t \in [0, T] \right\}$$

and

$$\mathcal{N}_t := \left\{ L_t = Y_t - X_{t+\Delta}^{\mathbf{u}} \mid \mathbf{u} \in \mathcal{A} \right\}.$$

Then the family $(\xi_t)_{t \in [0, T]}$ of maps ξ_t with

$$\xi_t: \mathcal{N}_t \rightarrow \mathcal{L}^1(\Omega, \mathcal{F}_t, \mathbb{P})$$

is called a dynamic risk measure. Common dynamic risk measures are presented in the next example.

Example 2.1.2.

- Given a time t and a probability level $\alpha \in (0, 1)$, the **Dynamic Value at Risk** denoted by VaR_t^α is the loss over $[t, t + \Delta]$ that is exceeded only with the (small) conditional probability α , thus

$$\xi_t(L_t) = \text{VaR}_t^\alpha(L_t) := \inf \{l \in \mathbb{R} \mid \mathbb{P}(L_t > l \mid \mathcal{F}_t) \leq \alpha\}.$$

- Given a time t and a probability level $\alpha \in (0, 1)$, the **Dynamic Tail Conditional Expectation** denoted by TCE_t^α is the conditional expected value of the loss exceeding VaR_t^α , thus

$$\xi_t(L_t) = \text{TCE}_t^\alpha(L_t) := \mathbb{E}_t [L_t \mid L_t > \text{VaR}_t^\alpha(L_t)],$$

where $\mathbb{E}_t[\cdot]$ denotes the conditional expectation given the information known up to time t .

- Given a time t , the **Dynamic Expected Loss** denoted by EL_t is the conditional expected value of “positive” losses, thus

$$\xi_t(L_t) = \text{EL}_t(L_t) := \mathbb{E}_t [L_t^+],$$

where $x^+ = \max(x, 0)$.

In the following we only consider those risk measures that can be written as $\xi_t(L_t) = \tilde{\xi}(t, X_t^u, \boldsymbol{\pi}_t, c_t)$ for all $t \in [0, T]$ and some measurable function $\tilde{\xi}: [0, T] \times \mathbb{R}_+ \times \mathcal{K} \rightarrow \mathbb{R}$. The risk measures presented in Example 2.1.2 belong to this class of risk measures. This becomes obvious if we recall that loss is defined by $L_t = Y_t - X_{t+\Delta}^u$ and that the benchmark is of the form $Y_t = \tilde{f}(t, X_t^u)$. Thus, we have to know the conditional distribution of $X_{t+\Delta}^u$ given X_t^u at any time t to explicitly compute the risk measures above. From Equation (2.2) and Equation (2.3) it is easily seen that the distribution of the investor’s wealth at a future date depends on the portfolio-proportion process $\mathbf{u} = (\boldsymbol{\pi}, c)$. For the purposes of risk measurement it is common practice to approximate this distribution (for elaborations see CUOCO ET AL. [24]). Let us consider the random variable

$$\mathcal{X} = \mathcal{X}(x, \bar{\boldsymbol{\pi}}, \bar{c}) = x \exp \left\{ \left(r + \bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) - \bar{c} - \frac{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2}{2} \right) \Delta + \bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}(\mathbf{W}_{t+\Delta} - \mathbf{W}_t) \right\},$$

where $\Delta > 0$, $x > 0$ and $(\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathcal{K}$ are given. It is easily seen that \mathcal{X} is log-normally distributed. More precisely, the law of $\mathcal{X}(x, \bar{\boldsymbol{\pi}}, \bar{c})$ is the one of xe^Z , where Z is a normally distributed random variable with mean $\left(r + \bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) - \bar{c} - \frac{1}{2}\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2 \right) \Delta$ and variance $\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2 \Delta$. We immediately obtain from Equation (2.3) that, given a

portfolio-proportion process $\mathbf{u} = (\boldsymbol{\pi}, c)$ and the associated portfolio wealth $X_t^{\mathbf{u}}$ at time t , the random variable $\mathcal{X}(X_t^{\mathbf{u}}, \boldsymbol{\pi}_t, c_t)$ is the value of the portfolio wealth at time $t + \Delta$, if the portfolio-proportion process is kept constant at $(\bar{\boldsymbol{\pi}}, \bar{c}) = (\boldsymbol{\pi}_t, c_t)$ between time t and $t + \Delta$. Then, $X_{t+\Delta}^{\mathbf{u}}$ is - conditionally on \mathcal{F}_t - distributed as $\mathcal{X}(X_t^{\mathbf{u}}, \boldsymbol{\pi}_t, c_t)$. Using this approximation we obtain the following formulas for the risk measures introduced in the example above.

Lemma 2.1.3.

1. *The Dynamic Value at Risk at time t can be written as*

$$\text{VaR}_t^\alpha(L_t) = \tilde{\xi}(t, X_t^{\mathbf{u}}, \boldsymbol{\pi}_t, c_t),$$

where

$$\begin{aligned} \tilde{\xi}(t, x, \bar{\boldsymbol{\pi}}, \bar{c}) = \tilde{f}(t, x) - x \exp \left[\left(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} - \frac{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2}{2} \right) \Delta \right. \\ \left. + \Phi^{-1}(\alpha) \|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\| \sqrt{\Delta} \right]. \end{aligned}$$

Here $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ denote the normal distribution and the inverse distribution functions, respectively.

2. *The Dynamic Tail Conditional Expectation at time t can be written as*

$$\text{TCE}_t^\alpha(L_t) = \tilde{\xi}(t, X_t^{\mathbf{u}}, \boldsymbol{\pi}_t, c_t),$$

where

$$\begin{aligned} \tilde{\xi}(t, x, \bar{\boldsymbol{\pi}}, \bar{c}) = \tilde{f}(t, x) - \frac{x}{\alpha} \left[\exp \{ (\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c}) \Delta \} \right. \\ \left. \Phi(\Phi^{-1}(\alpha) - \|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\| \sqrt{\Delta}) \right]. \end{aligned}$$

3. *The Dynamic Expected Loss at time t can be written as*

$$\text{EL}_t(L_t) = \tilde{\xi}(t, X_t^{\mathbf{u}}, \boldsymbol{\pi}_t, c_t),$$

where

$$\tilde{\xi}(t, x, \bar{\boldsymbol{\pi}}, \bar{c}) = \tilde{f}(t, x) \Phi(d_1) - x \exp \{ (\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c}) \Delta \} \Phi(d_2)$$

and

$$d_{1/2} = \frac{1}{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\| \sqrt{\Delta}} \left[\ln \left(\frac{\tilde{f}(t, x)}{x} \right) - \left(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} \mp \frac{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2}{2} \right) \Delta \right].$$

Proof. The first and the second claim are proved in [1, pp. 144-145] and [1, pp. 145-147], respectively. The proof of the third claim is presented in Section 2.4.1. \square

A risk constraint is imposed on the portfolio-proportion process by requiring that $\mathbf{u}_t = (\boldsymbol{\pi}_t, c_t)$ takes values in the set $\mathcal{K}^R(t, X_t^{\mathbf{u}})$ at any time t , where $\mathcal{K}^R(t, X_t^{\mathbf{u}})$ is defined by

$$\mathcal{K}^R(t, x) = \{(\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathbb{R}^d \times [0, \infty) \mid \tilde{\xi}(t, x, \bar{\boldsymbol{\pi}}, \bar{c}) \leq \tilde{\varepsilon}(t, x)\}, \quad x > 0.$$

Here, $\tilde{\varepsilon}: [0, T] \times \mathbb{R}_+ \rightarrow [0, \infty)$ is a measurable function which represents the bound on the risk constraint and may depend on time and wealth. Then the set of admissible strategies which continuously satisfy the imposed risk constraint reads as

$$\mathcal{A}^R := \{(\mathbf{u}_t)_{t \in [0, T]} \in \mathcal{A} \mid \mathbf{u}_t \in \mathcal{K}^R(t, X_t^{\mathbf{u}}) \text{ for all } t \in [0, T]\}.$$

We assume that the benchmark $Y_t = \tilde{f}(t, X_t^{\mathbf{u}})$ and the bound $\tilde{\varepsilon}(t, X_t^{\mathbf{u}})$ in the definition of the risk constraint are chosen such that $\mathcal{A}^R \neq \emptyset$ is satisfied, i.e., there exist admissible strategies. Such problems are examined in GABIH ET AL. [35], where the shortfall risk is measured in terms of the expected loss and applied in a static manner.

2.1.3 Optimization under risk constraints

In this section we introduce portfolio optimization problems under dynamic risk constraints. The investor derives utility from intermediate consumption and from terminal wealth while a risk constraint is imposed that has to be satisfied. The investor's performance criterion is the expected value of the utility of intermediate consumption and terminal wealth. Thus, the objective is to maximize

$$\mathbb{E}_{0, x_0} \left[\int_0^T U_1(C_t) dt + U_2(X_T^{\mathbf{u}}) \right]$$

over all portfolio-proportion processes $\mathbf{u} = (\boldsymbol{\pi}, c) \in \mathcal{A}^R$. Note that $C_t = c_t X_t^{\mathbf{u}}$. Here, $U_1: [0, \infty) \rightarrow \mathbb{R} \cup \{-\infty\}$ and $U_2: [0, \infty) \rightarrow \mathbb{R} \cup \{-\infty\}$ denote (time-independent) utility functions (i.e., U_1 and U_2 are strictly increasing, strictly concave and twice continuously differentiable on $(0, \infty)$). The term $\mathbb{E}_{t, x}[\cdot]$ denotes the conditional expectation given the information known up to time t and $X_t = x$.

Remark 2.1.4. We consider time-independent utility functions for the sake of notational simplicity but time-dependent utility functions can be treated in the same way. A typical example of a time-dependent utility function $U: [0, T] \times [0, \infty) \rightarrow \mathbb{R} \cup \{-\infty\}$ is

$$U(t, x) = e^{-\theta t} U_1(x), \quad \theta > 0.$$

Thus, discounting might be included.

We tackle the optimization problem by using dynamic programming techniques and embed it into a family of optimization problems. Given a portfolio-proportion process $\mathbf{u} = (\mathbf{u}_t)_{t \in [0, T]} \in \mathcal{A}^R$ we define the reward function $\mathcal{J}(\cdot, \cdot, \mathbf{u})$ for all $(t, x) \in [0, T] \times \mathbb{R}_+$ by

$$\mathcal{J}(t, x, \mathbf{u}) = \mathbb{E}_{t, x} \left[\int_t^T U_1(c_s X_s^{\mathbf{u}}) ds + U_2(X_T^{\mathbf{u}}) \right].$$

The objective is to maximize the reward function over all admissible processes which continuously satisfy the imposed risk constraint. We define the value function V for all $(t, x) \in [0, T] \times \mathbb{R}_+$ by

$$V(t, x) = \sup_{\mathbf{u} \in \mathcal{A}^R} \mathcal{J}(t, x, \mathbf{u}). \quad (2.4)$$

A portfolio-proportion process \mathbf{u}^* with $V = \mathcal{J}(\cdot, \cdot, \mathbf{u}^*)$ is called optimal. The Hamilton-Jacobi-Bellman (HJB) equation for V is derived by applying the dynamic programming principle (see Section 2.5.1) which yields

$$\frac{\partial}{\partial t} V(t, x) + \sup_{\bar{\mathbf{u}} = (\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathcal{K}^R(t, x)} \{U_1(\bar{c}x) + \mathcal{H}^{\bar{\mathbf{u}}} V(t, x)\} = 0 \quad (2.5)$$

for $(t, x) \in [0, T] \times \mathbb{R}_+$ with terminal condition $V(T, x) = U_2(x)$ for $x \in \mathbb{R}_+$. Here, the operator $\mathcal{H}^{\bar{\mathbf{u}}}$ acting on $\mathcal{C}^{1,2}([0, T] \times \mathbb{R}_+)$ is defined by

$$\mathcal{H}^{\bar{\mathbf{u}}} = [x(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c})] \frac{\partial}{\partial x} + \frac{1}{2} x^2 \bar{\boldsymbol{\pi}}' \boldsymbol{\sigma} \boldsymbol{\sigma}' \bar{\boldsymbol{\pi}} \frac{\partial^2}{\partial x^2}.$$

Note that $\mathcal{H}^{\bar{\mathbf{u}}}$ is the generator of the controlled wealth process when the portfolio-proportion process takes the value $\bar{\mathbf{u}} = (\bar{\boldsymbol{\pi}}, \bar{c})$. Unlike many stochastic control problems our formulation poses a state-dependent set of admissible controls. This difficulty can be handled by embedding the state-dependent set $\mathcal{K}^R(t, X_t^{\mathbf{u}})$ into a compact set $\bar{\mathcal{K}}$ (see SHARDIN & WUNDERLICH [75]). Using the techniques presented in FLEMING & SONER [33, Chapter III, Theorem 8.1] one can show that if there exists a classical solution W of the HJB equation (2.5), then W coincides with the value function V of the control problem (2.4). Furthermore, if there exists a measurable function $\tilde{\mathbf{u}}^*: [0, T] \times \mathbb{R}_+ \rightarrow \mathcal{K}$ satisfying $\tilde{\mathbf{u}}^*(t, x) \in \mathcal{K}^R(t, x)$ for all $(t, x) \in [0, T] \times \mathbb{R}_+$ and such that for every $(t, x) \in [0, T] \times \mathbb{R}_+$ the value of $\tilde{\mathbf{u}}^*$ at (t, x) is the unique maximizer of the problem

$$\sup_{\bar{\mathbf{u}} = (\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathcal{K}^R(t, x)} \{U_1(\bar{c}x) + \mathcal{H}^{\bar{\mathbf{u}}} V(t, x)\}$$

then one can show that the optimal portfolio-proportion process is given by $\mathbf{u}_t^* = \tilde{\mathbf{u}}^*(t, X_t^{u^*})$ and satisfies $\mathbf{u}^* \in \mathcal{A}^R$.

For the case of power utility, i.e.,

$$U_1(x) = U_2(x) = \frac{x^{1-\gamma}}{1-\gamma}, \quad 0 < \gamma < 1,$$

there is a closed-form expression for the optimal portfolio-proportion process when no risk constraint is imposed, see KORN [50]. We call this process the Merton portfolio-proportion process and it is given by

$$\begin{aligned} \boldsymbol{\pi}^M(t, x) &= \boldsymbol{\pi}^M = (\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - \mathbf{1}r) \frac{1}{\gamma} \quad \text{and} \\ c^M(t, x) &= c^M(t) = (\tau^{-1} + (1 - \tau^{-1})e^{-\tau(T-t)})^{-1}, \end{aligned} \tag{2.6}$$

where

$$\tau = \left[- (1 - \gamma) \left(\frac{(\boldsymbol{\mu} - \mathbf{1}r)'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - \mathbf{1}r)}{2\gamma} + r \right) \right] \frac{1}{\gamma}.$$

An investor using the Merton portfolio-proportion process is called a Merton investor and the associated value function is denoted by $V^M(t, x)$. We write X^M when the wealth is controlled by the Merton portfolio-proportion process $\mathbf{u}^M := (\boldsymbol{\pi}^M, c^M)$.

2.2 Discrete-time optimization

2.2.1 Model

In this section we suppose that the trading interval $[0, T]$ is divided into N periods of length Δ and trading only takes place at the beginning of each of the N periods. The trading times are denoted by $t_n := n\Delta$, $n = 0, \dots, N-1$, and for the time horizon T we write $t_N := T = N\Delta$. Uncertainty is modeled by the filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ where the filtration $\mathbb{G} = (\mathcal{G}_{t_n})_{n=0, \dots, N}$ is generated by the (discretized) m -dimensional Brownian motion $(\mathbf{W}_{t_n})_{n \in \{0, \dots, N\}}$ and augmented by all the \mathbb{P} -null sets of Ω . In what follows we will consider an N -period financial market which results from a discretization of the Black-Scholes-Merton model considered in the continuous-time case and consists of one risk-free and d risky securities. Thus, the price process of the bond is given by $S_{t_0}^0 \equiv 1$ and

$$S_{t_{n+1}}^0 = S_{t_n}^0 e^{r\Delta}, \quad n = 0, \dots, N-1.$$

Recall, $r \geq 0$ denotes the continuously compounded interest rate. For all $k = 1, \dots, d$ the price process of the k -th risky security is given by $S_{t_0}^k = s^k$ and

$$S_{t_{n+1}}^k = S_{t_n}^k \tilde{R}_{t_{n+1}}^k, \quad n = 0, \dots, N-1,$$

where $(\tilde{\mathbf{R}}_{t_{n+1}})_{n \in \{0, \dots, N-1\}} = (\tilde{R}_{t_{n+1}}^1, \tilde{R}_{t_{n+1}}^2, \dots, \tilde{R}_{t_{n+1}}^d)_{n \in \{0, \dots, N-1\}}$ denotes the return process and $\tilde{R}_{t_{n+1}}^k$ is defined by

$$\tilde{R}_{t_{n+1}}^k = \exp \left\{ \left(\mu^k - \frac{1}{2} \sum_{j=1}^m (\sigma^{kj})^2 \right) \Delta + \sum_{j=1}^m \sigma^{kj} (W_{t_{n+1}}^j - W_{t_n}^j) \right\}$$

for $k = 1, \dots, d$ and $n = 0, \dots, N-1$. The random variables $\tilde{R}_{t_{n+1}}^k$ are log-normally distributed and represent the relative price change of the risky securities in the time interval $[t_n, t_{n+1})$.

Analogously to the continuous-time case the investor starts with an initial wealth $x_0 > 0$ and is allowed to invest in the financial market and to consume the wealth. In contrast to the continuous-time case the investor can adjust only at the beginning of each of the N periods the amount of wealth invested into the financial market and the amount of wealth consumed. The amount invested in the d risky securities is denoted by $\varphi = (\varphi_{t_n})_{n \in \{0, \dots, N-1\}}$ and the amount which is consumed by $\eta = (\eta_{t_n})_{n \in \{0, \dots, N-1\}}$. Given an investment-consumption strategy $\nu = (\varphi, \eta)$ the associated wealth process evolves as follows

$$\begin{aligned} X_{t_{n+1}} &= e^{r\Delta} (X_{t_n} - \eta_{t_n} - \varphi'_{t_n} \mathbf{1}) + \varphi'_{t_n} \cdot \tilde{\mathbf{R}}_{t_{n+1}} \\ &= e^{r\Delta} (X_{t_n} - \eta_{t_n} + \varphi'_{t_n} \cdot \mathbf{R}_{t_{n+1}}), \end{aligned}$$

where $(\mathbf{R}_{t_{n+1}})_{n \in \{0, \dots, N-1\}} = (R_{t_{n+1}}^1, \dots, R_{t_{n+1}}^d)_{n \in \{0, \dots, N-1\}}$ denotes the relative discounted return process and $R_{t_{n+1}}^k$ is defined by

$$R_{t_{n+1}}^k = e^{-r\Delta} \tilde{R}_{t_{n+1}}^k - 1$$

for $k = 1, \dots, d$ and $n = 0, \dots, N-1$. The investment strategy (φ_{t_n}) and the consumption strategy (η_{t_n}) are assumed to be \mathbb{G} -adapted. Moreover we restrict to strategies which achieve a positive wealth for all N periods. Thus, in contrast to the continuous-time case, for an investor who trades only at discrete points in time it is not admissible to sell stocks short or to take out a loan to buy stocks because in a time interval of length Δ the stock price could move unfavorably for the investor, and the wealth would become negative. Therefore, at any trading time t_n , $n \in \{0, \dots, N-1\}$, the investor must choose the amount φ_{t_n} of current wealth X_{t_n} that is invested in the stocks and the amount η_{t_n} of current wealth X_{t_n} that is consumed in such a way that $\nu_{t_n} = (\varphi_{t_n}, \eta_{t_n}) \in \mathcal{K}_D(X_{t_n})$ holds, where

$$\mathcal{K}_D(x) = \{ \bar{\nu} = (\bar{\varphi}, \bar{\eta}) \mid 0 \leq \bar{\eta} \leq x \text{ and } 0 \leq \bar{\varphi}' \mathbf{1} \leq x - \bar{\eta} \}, \quad x > 0.$$

Furthermore, we only consider strategies which are of Markov type, i.e.,

$$\nu_{t_n} = \tilde{\nu}(t_n, X_{t_n}) \quad \text{for all } n = 0, \dots, N-1,$$

where $\tilde{\nu}: [0, T] \times \mathbb{R}_+ \rightarrow [0, \infty)^{d+1}$ is a measurable function. Such strategies are called admissible and the set of admissible strategies is denoted by \mathcal{A}_Δ^0 , thus

$$\mathcal{A}_\Delta^0 := \left\{ (\boldsymbol{\nu}_{t_n})_{n \in \{0, \dots, N-1\}} \mid \boldsymbol{\nu} \text{ is } \mathbb{G}\text{-adapted, } \boldsymbol{\nu}_{t_n} = (\varphi_{t_n}, \eta_{t_n}) \in \mathcal{K}_D(X_{t_n}), \right. \\ \left. \boldsymbol{\nu}_{t_n} = \tilde{\nu}(t_n, X_{t_n}) \text{ and } X_{t_n} > 0 \text{ for all } n \in \{0, \dots, N-1\} \right\}.$$

It is shown in HINDERER [42, Theorem 18.4] that more general strategies which depend on the complete history of the process instead of being Markovian ones, do not increase the value of the maximization problem that we will consider later on. We write in the following X^ν instead of X to emphasize that the wealth is controlled by the investment-consumption strategy $\boldsymbol{\nu} = (\varphi, \eta)$.

2.2.2 Dynamic risk constraints

In this section we introduce dynamic risk measures in a discrete-time setting. In an analogous manner to the continuous-time case the loss over the period $[t_n, t_{n+1})$ is defined by $L_{t_n} := Y_{t_n} - X_{t_{n+1}}^\nu$ with Y_{t_n} being a \mathcal{G}_{t_n} -measurable benchmark (see Example 2.1.1) prescribed at time t_n . Let $\mathcal{N}_{t_n}^\Delta := \{L_{t_n} \mid \boldsymbol{\nu} \in \mathcal{A}_\Delta\}$, where

$$\mathcal{A}_\Delta := \left\{ (\boldsymbol{\nu}_{t_n})_{n \in \{0, \dots, N-1\}} \in \mathcal{A}_\Delta^0 \mid \mathbb{E}[|L_{t_n}|] < \infty \text{ for all } n \in \{0, \dots, N-1\} \right\}.$$

A dynamic risk measure (in discrete time) $(\psi_{t_n})_{n \in \{0, \dots, N-1\}}$ is a family of maps ψ_{t_n} with

$$\psi_{t_n}: \mathcal{N}_{t_n}^\Delta \rightarrow \mathcal{L}^1(\Omega, \mathcal{G}_{t_n}, \mathbb{P}).$$

We restrict to risk measures of the form $\psi_{t_n}(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}^\nu, \varphi_{t_n}, \eta_{t_n})$ for all $n \in \{0, \dots, N-1\}$ and some measurable function $\tilde{\psi}: [0, T] \times \mathbb{R}_+ \times [0, \infty)^{d+1} \rightarrow \mathbb{R}$. Lemma 2.2.1 below shows how the dynamic risk measures VaR, TCE, and EL can be computed explicitly in the discrete-time case if we consider a market with a single stock. In this case we can benefit from the fact that the wealth $X_{t_{n+1}}^\nu$ at time t_{n+1} is - conditionally on \mathcal{G}_{t_n} - (shifted) log-normally distributed. In the case $d > 1$ the wealth $X_{t_{n+1}}^\nu$ at time t_{n+1} given \mathcal{G}_{t_n} is a sum of dependent log-normally distributed random variables and closed-form expressions for these risk measures are not available.

Lemma 2.2.1. *The following statements hold.*

1. Given a probability level $\alpha \in (0, 1)$, the Dynamic Value at Risk at time t_n , $n = 0, \dots, N-1$ can be written as $\text{VaR}_{t_n}^\alpha(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}^\nu, \varphi_{t_n}, \eta_{t_n})$, where

$$\tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) = \tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) - \exp \left\{ \Phi^{-1}(\alpha) \sigma \sqrt{\Delta} + \left(\mu - \frac{\sigma^2}{2} \right) \Delta \right\} \bar{\varphi}.$$

Recall that $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ denote the cumulative distribution function of the standard normal distribution and its inverse, respectively.

2. Given a probability level $\alpha \in (0, 1)$, the Dynamic Tail Conditional Expectation at time t_n , $n = 0, \dots, N-1$ can be written as $TCE_{t_n}^\alpha(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}^\nu, \varphi_{t_n}, \eta_{t_n})$, where

$$\tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) = \tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) - \frac{1}{\alpha} e^{\mu\Delta} \Phi\left(\Phi^{-1}(\alpha) - \sigma\sqrt{\Delta}\right) \bar{\varphi}.$$

3. The Dynamic Expected Loss at time t_n , $n = 0, \dots, N-1$ can be written as $EL_{t_n}(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}^\nu, \varphi_{t_n}, \eta_{t_n})$, where

$$\begin{aligned} \tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) &= \left(\tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) \right) \Phi(d_1) - e^{\mu\Delta} \Phi(d_2) \bar{\varphi}, \\ d_{1/2} &= \frac{1}{\sigma\sqrt{\Delta}} \left[\ln\left(\frac{\tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi})}{\bar{\varphi}}\right) - (\mu \mp \sigma^2/2)\Delta \right]. \end{aligned}$$

Proof. The proof is presented in Section 2.4.2. \square

A risk constraint is imposed on the strategy by requiring that at the beginning of each period $[t_n, t_{n+1})$, $n \in \{0, \dots, N-1\}$, the investor must decide how much of the wealth is invested in stocks (φ_{t_n}) and how much is consumed (η_{t_n}) such that $\nu_{t_n} = (\varphi_{t_n}, \eta_{t_n}) \in \mathcal{K}_D^R(t_n, X_{t_n}^\nu)$ with

$$\mathcal{K}_D^R(t_n, x) = \left\{ \bar{\nu} = (\bar{\varphi}, \bar{\eta}) \in \mathcal{K}_D(x) \mid \tilde{\psi}(t_n, x, \bar{\varphi}, \bar{\eta}) \leq \tilde{\epsilon}(t_n, x) \right\}, \quad x > 0.$$

Here, $\tilde{\epsilon}: [0, T] \times \mathbb{R}_+ \rightarrow [0, \infty)$ is a measurable function which represents the bound on the risk constraint and may depend on time and wealth. The set of admissible strategies reads as

$$\mathcal{A}_D^R := \left\{ (\nu_{t_n})_{n \in \{0, \dots, N-1\}} \in \mathcal{A}_\Delta \mid \nu_{t_n} \in \mathcal{K}_\Delta^R(t_n, X_{t_n}^\nu), n \in \{0, \dots, N-1\} \right\}.$$

We assume that the benchmark $\tilde{f}(t_n, X_{t_n}^\nu)$ and bound $\tilde{\epsilon}(t_n, X_{t_n}^\nu)$ are specified in a way that $\mathcal{A}_D^R \neq \emptyset$ is satisfied, i.e., there exist admissible strategies.

2.2.3 Optimization under risk constraints

In this section we formulate the investor's portfolio optimization problem under a dynamic risk constraint in a discrete-time setting and present a backward induction algorithm for solving the problem. Given an initial wealth $X_0 = x_0$ the investor's investment-consumption problem is to decide how much wealth is invested in stocks and how much is consumed so that the expected value of the utility from consumption and terminal wealth,

$$\mathbb{E}_{t_0, x_0} \left[\sum_{n=0}^{N-1} U_1(\eta_{t_n}) + U_2(X_{t_N}^\nu) \right],$$

is maximized over all strategies $\boldsymbol{\nu} = (\boldsymbol{\varphi}, \eta) \in \mathcal{A}_\Delta^R$. This problem can be solved by using the theory of Markov Decision Processes (see Section 2.5.2). The value function V is defined by

$$V(t_n, x) = \sup_{(\boldsymbol{\varphi}, \eta) \in \mathcal{A}_\Delta^R} \mathbb{E}_{t_n, x} \left[\sum_{k=n}^{N-1} U_1(\eta_{t_k}) + U_2(X_{t_N}^\nu) \right]$$

and the sequence $(V(t_n, x))_{n=0, \dots, N-1}$ can be computed by solving the optimality equation

$$\begin{aligned} V(t_N, x) &= U_2(x), \\ V(t_n, x) &= \sup_{(\bar{\boldsymbol{\varphi}}, \bar{\eta}) \in \mathcal{K}_D^R(t_n, x)} \left\{ U_1(\bar{\eta}) + \mathbb{E}_{t_n, x} \left[V(t_{n+1}, e^{r\Delta} (x - \bar{\eta} + \bar{\boldsymbol{\varphi}}' \cdot \mathbf{R}_{t_{n+1}})) \right] \right\} \quad (2.7) \\ & \quad n = N - 1, \dots, 0. \end{aligned}$$

The optimal strategy $\boldsymbol{\nu}^* = (\boldsymbol{\varphi}^*, \eta^*)$ is generated by the sequence of maximizers of $V(t_1, x), \dots, V(t_N, x)$.

For the case of power utility, i.e.,

$$U_1(x) = U_2(x) = \frac{x^{1-\gamma}}{1-\gamma}, \quad \gamma \in (0, 1),$$

we can specify the recursion above. Let ζ and $\boldsymbol{\beta}$ denote the consumption and investment proportion respectively, i.e.,

$$\zeta_{t_n} := \frac{\eta_{t_n}}{X_{t_n}^\nu} \quad \text{and} \quad \beta_{t_n}^i := \frac{\varphi_{t_n}^i}{X_{t_n}^\nu - \eta_{t_n}}, \quad \text{for } n = 0, \dots, N-1, i = 1, \dots, d.$$

Note that $\zeta_{t_n} \in [0, 1]$ and $\boldsymbol{\beta}_{t_n} \in \mathcal{P}$ where $\mathcal{P} := \{\mathbf{p} \in [0, 1]^d \mid p^1 + \dots + p^d \leq 1\}$ denotes the simplex in \mathbb{R}^d . In order to use proportions instead of amounts of consumption and investment we redefine $\tilde{\psi}$ by

$$\tilde{\psi}_{rel}(t_n, X_{t_n}^\nu, \boldsymbol{\beta}_{t_n}, \zeta_{t_n}) = \tilde{\psi}(t_n, X_{t_n}^\nu, (1 - \zeta_{t_n})X_{t_n}^\nu \boldsymbol{\beta}_{t_n}, \zeta_{t_n} X_{t_n}^\nu).$$

Then the risk measures can be written as $\psi_{t_n}(L_{t_n}) = \tilde{\psi}_{rel}(t_n, X_{t_n}^\nu, \boldsymbol{\beta}_{t_n}, \zeta_{t_n})$ for all $n \in \{0, \dots, N-1\}$ and we obtain by the optimality equation (2.7) the following backward recursion.

Theorem 2.2.2. *The value function for $n = 0, \dots, N$ takes the form*

$$V(t_n, x) = \frac{x^{1-\gamma}}{1-\gamma} \cdot d_{t_n}, \quad x > 0,$$

where (d_{t_n}) satisfies the backward recursion

$$\begin{aligned} d_{t_N} &= 1, \\ d_{t_n} &= \sup_{0 \leq \bar{\zeta} \leq 1} \left\{ \bar{\zeta}^{1-\gamma} + (1 - \bar{\zeta})^{1-\gamma} e^{r\Delta(1-\gamma)} \right. \\ &\quad \left. \times \left(\sup_{\bar{\beta} \in \mathcal{B}(t_n, x, \bar{\zeta})} \mathbb{E}[(1 + \bar{\beta}' \cdot \mathbf{R}_{t_{n+1}})^{1-\gamma}] \right) d_{t_{n+1}} \right\}, \quad n = N-1, \dots, 0 \end{aligned} \quad (2.8)$$

with

$$\mathcal{B}(t_n, x, \bar{\zeta}) = \{\bar{\beta} \in \mathcal{P} \mid \tilde{\psi}_{rel}(t_n, x, \bar{\beta}, \bar{\zeta}) \leq \tilde{\epsilon}(t_n, x)\}.$$

Proof. The proof is given by mathematical induction. The basic step is to show that the statement

$$V(t_N, x) = \frac{x^{1-\gamma}}{1-\gamma} \cdot d_{t_N}, \quad d_{t_N} = 1$$

holds. This follows straightforwardly from the optimality equation (2.7). In the inductive step we assume that the statement

$$V(t_n, x) = \frac{x^{1-\gamma}}{1-\gamma} \cdot d_{t_n}$$

holds for some $n \in \{1, \dots, N\}$ and show that it also holds for $n-1$. This can be done as follows. From the optimality equation (2.7) we obtain

$$\begin{aligned} V(t_{n-1}, x) &= \sup_{(\bar{\varphi}, \bar{\eta}) \in \mathcal{K}_D^R(t_{n-1}, x)} \left\{ \mathbb{E}_{t_{n-1}, x} [V(t_n, e^{r\Delta}(x - \bar{\eta} + \bar{\varphi}' \cdot \mathbf{R}_{t_n}))] \right. \\ &\quad \left. + \frac{\bar{\eta}^{1-\gamma}}{1-\gamma} \right\}. \end{aligned}$$

Using the inductive hypothesis, the right-hand side can be rewritten as

$$\sup_{(\bar{\varphi}, \bar{\eta}) \in \mathcal{K}_D^R(t_{n-1}, x)} \left\{ \mathbb{E}_{t_{n-1}, x} \left[\frac{1}{1-\gamma} (e^{r\Delta}(x - \bar{\eta} + \bar{\varphi}' \cdot \mathbf{R}_{t_n}))^{1-\gamma} d_{t_n} \right] + \frac{\bar{\eta}^{1-\gamma}}{1-\gamma} \right\}.$$

Substituting $\bar{\eta} = \bar{\zeta}x$ and $\bar{\varphi} = (1 - \bar{\zeta})x\bar{\beta}$ we find

$$\begin{aligned}
& \sup_{\substack{0 \leq \bar{\zeta} \leq 1 \\ \bar{\beta} \in \mathcal{B}(t_{n-1}, x, \bar{\zeta})}} \left\{ \mathbb{E}_{t_{n-1}, x} \left[\frac{1}{1 - \gamma} \left(e^{r\Delta} \left(x - \bar{\zeta}x + x(1 - \bar{\zeta}) \cdot \bar{\beta}' \mathbf{R}_{t_n} \right) \right)^{1-\gamma} d_{t_n} \right] \right. \\
& \quad \left. + \frac{(\bar{\zeta}x)^{1-\gamma}}{1 - \gamma} \right\} \\
= & \sup_{\substack{0 \leq \bar{\zeta} \leq 1 \\ \bar{\beta} \in \mathcal{B}(t_{n-1}, x, \bar{\zeta})}} \left\{ \frac{x^{1-\gamma}}{1 - \gamma} e^{r\Delta(1-\gamma)} (1 - \bar{\zeta})^{1-\gamma} \mathbb{E} \left[(1 + \bar{\beta}' \mathbf{R}_{t_n})^{1-\gamma} \right] d_{t_n} \right. \\
& \quad \left. + \frac{x^{1-\gamma}}{1 - \gamma} \bar{\zeta}^{1-\gamma} \right\} \\
= & \frac{x^{1-\gamma}}{1 - \gamma} \sup_{0 \leq \bar{\zeta} \leq 1} \left\{ \bar{\zeta}^{1-\gamma} + e^{r\Delta(1-\gamma)} (1 - \bar{\zeta})^{1-\gamma} \right. \\
& \quad \left. \times \sup_{\bar{\beta} \in \mathcal{B}(t_{n-1}, x, \bar{\zeta})} \mathbb{E} \left[(1 + \bar{\beta}' \mathbf{R}_{t_n})^{1-\gamma} \right] \cdot d_{t_n} \right\}.
\end{aligned}$$

This shows that indeed it holds

$$V(t_{n-1}, x) = \frac{x^{1-\gamma}}{1 - \gamma} \cdot d_{t_{n-1}}.$$

In the second equation we have used that the relative discounted return \mathbf{R}_{t_n} is independent of $\mathcal{G}_{t_{n-1}}$ and in the third equation we used that $e^{r\Delta(1-\gamma)}(1 - \bar{\zeta})^{1-\gamma}$ is non-negative. Hence, the supremum over $(\bar{\beta}, \bar{\zeta})$ can be obtained by the iterated supremum as given. The proof is complete by mathematical induction. \square

Note that in contrast to the continuous-time case, there is no closed-form solution to the unconstrained problem in discrete time. The Merton portfolio-proportion strategy $(\beta_{t_n}^M, \zeta_{t_n}^M)_{n=0, \dots, N-1}$ also has to be computed by backward recursion as in the above theorem where the set \mathcal{B} is replaced by the simplex \mathcal{P} (see BÄUERLE & RIEDER [7]). Then, we obtain

$$\beta_{t_n}^M = \arg \max_{\beta \in \mathcal{P}} \mathbb{E} \left[(1 + \beta' \mathbf{R}_{t_{n+1}})^{1-\gamma} \right] \quad \text{and} \quad \zeta_{t_n}^M = \left[1 + (e^{r\Delta(1-\gamma)} v_{t_n} d_{t_n})^{\frac{1}{\gamma}} \right]^{-1}$$

for $n = 0, \dots, N - 1$ with

$$v_{t_n} := \sup_{\beta \in \mathcal{P}} \mathbb{E} \left[(1 + \beta' \cdot \mathbf{R}_{t_{n+1}})^{1-\gamma} \right], \quad n = 0, \dots, N - 1.$$

The value function and wealth obtained by an investor using the Merton portfolio-proportion strategy are denoted by V^M and X^M , respectively.

2.3 Numerical examples for power utility

In this section we solve the continuous-time optimization problem

$$\sup_{(\boldsymbol{\pi}, c) \in \mathcal{A}^R} \mathbb{E}_{0, x_0} \left[\int_0^T U_1(C_t) dt + U_2 \left(X_T^{(\boldsymbol{\pi}, c)} \right) \right]$$

and the discrete-time optimization problem

$$\sup_{(\boldsymbol{\varphi}, \eta) \in \mathcal{A}_D^R} \mathbb{E}_{0, x_0} \left[\sum_{n=0}^{N-1} U_1(\eta_{t_n}) + U_2 \left(X_{t_N}^{(\boldsymbol{\varphi}, \eta)} \right) \right]$$

numerically for the case of power utility and compare the solutions. In Section 2.3.1 and Section 2.3.2 a numerical approach is presented for solving the continuous-time and discrete-time optimization problem, respectively. Examples based on a financial market consisting of one risk-free and one risky security are considered in Section 2.3.3, whereas the financial market considered in Section 2.3.4 consists of one risk-free and two risky securities.

2.3.1 Continuous-time optimization problem

We start with the continuous-time problem by numerically solving the HJB equation

$$\frac{\partial}{\partial t} V(t, x) + \sup_{\bar{\boldsymbol{u}} = (\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathcal{K}^R(t, x)} \{U_1(\bar{c}x) + \mathcal{H}^{\bar{\boldsymbol{u}}} V(t, x)\} = 0,$$

for $(t, x) \in [0, T) \times \mathbb{R}_+$ with terminal condition $V(T, x) = U_2(x)$ for $x \in \mathbb{R}_+$. An approximation of the value function and of the optimal portfolio-proportion process is obtained by applying the following policy improvement (PI) algorithm (for a proof see BÄUERLE & RIEDER [8]).

Algorithm 2.3.1 (Policy Improvement).

1. Let $k = 0$. Choose an arbitrary admissible initial strategy $\boldsymbol{u}^0 = (\boldsymbol{\pi}^0, c^0) \in \mathcal{A}^R$.
2. Compute the associated reward function for the fixed strategy $\bar{\boldsymbol{u}}^k = (\bar{\boldsymbol{\pi}}^k, \bar{c}^k)$ by solving the linear partial differential equation (PDE)

$$\begin{aligned} 0 = & \frac{\partial}{\partial t} J^k(t, x) + U_1(\bar{c}^k x) + \left[x \left(\bar{\boldsymbol{\pi}}^{k'} (\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c}^k \right) \right] \frac{\partial}{\partial x} J^k(t, x) \\ & + \frac{1}{2} x^2 \bar{\boldsymbol{\pi}}^{k'} \boldsymbol{\sigma} \boldsymbol{\sigma}' \bar{\boldsymbol{\pi}}^k \frac{\partial^2}{\partial x^2} J^k(t, x), \quad (t, x) \in [0, T) \times \mathbb{R}_+ \end{aligned} \quad (2.9)$$

with terminal condition

$$J^k(T, x) = U_2(x), \quad x \in \mathbb{R}_+. \quad (2.10)$$

3. Compute the succeeding strategy $\mathbf{u}^{k+1} = (\boldsymbol{\pi}^{k+1}, c^{k+1})$ by solving the optimization problem

$$\mathbf{u}^{k+1} = \arg \max_{\bar{\mathbf{u}} = (\bar{\boldsymbol{\pi}}, \bar{c}) \in \mathcal{K}^R(t, x)} \{U_1(\bar{c}x) + \mathcal{H}^{\bar{\mathbf{u}}} J^k(t, x)\}. \quad (2.11)$$

4. Let $k = k + 1$ and go back to step 2.

The PI algorithm provides a numerical solution of the HJB equation. In each iteration we have to solve a linear PDE and a constrained optimization problem. The linear PDE is solved by using meshless methods and the constrained nonlinear optimization problem by using sequential quadratic programming methods. Both methods are elaborated in the following subsections.

Meshless method

The meshless method is applied to numerically solve the linear PDE (2.9) for the case of power utility. For an overview of the meshless method we refer to LI & CHEN [56, Chapter 10] and DUFFY [29, Chapter 16], respectively.

First, we discretize equation (2.9) in t by applying the Crank-Nicolson method (see THOMAS [78]) while still keeping the variable x continuous. Then we approximate the solution of the resulting ordinary differential equation (ODE) by using a special class of functions, the so-called radial basis functions (RBFs, see CHEN, FU & CHEN [23]).

In particular, we partition the interval $[0, T]$ into n_t equal sub-intervals of length Δt and define the nodes

$$t^n := n \cdot \Delta t, \quad \text{where } n \in \{0, \dots, n_t\}.$$

A semi-discretization of equation (2.9) in time using Crank-Nicolson leads to the following system of ordinary differential equations (where the iteration number k is omitted to simplify notation):

$$\begin{aligned} 0 = & \frac{J^n(x) - J^{n+1}(x)}{\Delta t} \\ & + \frac{1}{2} \left[[(r + (\boldsymbol{\pi}^n(x))'(\boldsymbol{\mu} - \mathbf{1}r))x - c^n(x)] \frac{d}{dx} J^n(x) \right. \\ & \quad \left. + \frac{1}{2} (\boldsymbol{\pi}^n(x))' \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi}^n(x) x^2 \frac{d^2}{dx^2} J^n(x) + \frac{(c^n(x))^{1-\gamma}}{1-\gamma} \right] \\ & + \frac{1}{2} \left[[(r + (\boldsymbol{\pi}^{n+1}(x))'(\boldsymbol{\mu} - \mathbf{1}r))x - c^{n+1}(x)] \frac{d}{dx} J^{n+1}(x) \right. \\ & \quad \left. + \frac{1}{2} (\boldsymbol{\pi}^{n+1}(x))' \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi}^{n+1}(x) x^2 \frac{d^2}{dx^2} J^{n+1}(x) + \frac{(c^{n+1}(x))^{1-\gamma}}{1-\gamma} \right], \end{aligned}$$

where $J^n(x)$ denotes the approximate value of J at node t^n and wealth x , $\boldsymbol{\pi}^n(x) := \boldsymbol{\pi}(t^n, x)$ and $c^n(x) := c(t^n, x)$. Multiplying by Δt and rewriting yields

$$\begin{aligned} & J^{n+1}(x) - \frac{\Delta t}{2} \left[[(r + (\boldsymbol{\pi}^{n+1}(x))'(\boldsymbol{\mu} - \mathbf{1}r))x - c^{n+1}(x)] \frac{d}{dx} J^{n+1}(x) \right. \\ & \quad \left. + \frac{1}{2} (\boldsymbol{\pi}^{n+1}(x))' \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi}^{n+1}(x) x^2 \frac{d^2}{dx^2} J^{n+1}(x) + \frac{(c^{n+1}(x))^{1-\gamma}}{1-\gamma} \right] \\ = & J^n(x) + \frac{\Delta t}{2} \left[[(r + (\boldsymbol{\pi}^n(x))'(\boldsymbol{\mu} - \mathbf{1}r))x - c^n(x)] \frac{d}{dx} J^n(x) \right. \\ & \quad \left. + \frac{1}{2} (\boldsymbol{\pi}^n(x))' \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi}^n(x) x^2 \frac{d^2}{dx^2} J^n(x) + \frac{(c^n(x))^{1-\gamma}}{1-\gamma} \right] \end{aligned}$$

or in the differential operator form

$$H_-^{n+1}(x)J^{n+1}(x) - d^{n+1}(x) = H_+^n(x)J^n(x) + d^n(x), \quad (2.12)$$

where we define the operators as

$$H_{\pm}^n(x) = 1 \pm \frac{\Delta t}{2} F^n(x), \quad d^n(x) = \frac{\Delta t}{2} \frac{(c^n(x))^{1-\gamma}}{1-\gamma}$$

and

$$F^n(x) = [(r + (\boldsymbol{\pi}^n(x))'(\boldsymbol{\mu} - \mathbf{1}r))x - c^n(x)] \frac{d}{dx} + \frac{1}{2} (\boldsymbol{\pi}^n(x))' \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi}^n(x) x^2 \frac{d^2}{dx^2}.$$

We now assume the solution J of equation (2.9) can be approximated by a linear combination of radial basis functions

$$J(t, x) \approx \sum_{m=1}^M \lambda_m(t) \phi(r_m) \quad \text{bzw.} \quad J^n(x) \approx \sum_{m=1}^M \lambda_m^n \phi(r_m), \quad x \in \mathbb{R}_+ \quad (2.13)$$

where r_m denotes the Euclidean distance between point x and x_m , i.e.,

$$r_m := \sqrt{(x - x_m)^2},$$

$(x_m)_{m=1}^M$ are given collocation points, λ_m^n are time-dependent coefficients to be determined and ϕ is the radial basis function. There are a number of types of RBFs whose characteristic feature is that their values depend only on the distance from some center point [29]. Four often used RBFs are the multi-quadrics (MQ) function

$$\phi(r_m) = (r_m^2 + (\alpha_c d_c)^2)^q,$$

the Gaussian function

$$\phi(r_m) = \exp \left\{ -\alpha_c \left(\frac{r_m}{d_c} \right)^2 \right\},$$

the thin plate spline (TPS) function

$$\phi(r_m) = r_m^q,$$

and the logarithmic radial basis function

$$\phi(r_m) = r_m^q \log(r_m),$$

where q and α_c are shape parameters of the RBFs (see LIU & GU [57]). The coefficient d_c is a size coefficient, usually the average nodal spacing of the M collocation points (see BELINHA [10]). In utilizing RBFs, their shape parameters need to be determined, which can be done by numerical examinations for given types of problems (see, e.g., WANG & LIU [80]). Here, the shape parameters can be adjusted by numerical examination of the Merton investor's value function for which an explicit formula exists. Letting the value x in equation (2.13) be a specific collocation point leads to the expression:

$$J^n(x_i) \approx \sum_{m=1}^M \lambda_m^n \phi(r_{im}), \quad (2.14)$$

where $r_{im} := \sqrt{(x_i - x_m)^2}$.

The ODEs (2.12) are now approximated at each allocation point by inserting the above expression (2.14) into equation (2.12), yielding the identity

$$\sum_{m=1}^M \lambda_m^{n+1} H_-^{n+1}(x_i) \phi(r_{im}) - d^{n+1}(x_i) = \sum_{m=1}^M \lambda_m^n H_+^n(x_i) \phi(r_{im}) + d^n(x_i), \quad 1 \leq i \leq M \quad (2.15)$$

or in matrix form

$$\mathbf{B}^{n+1} = \mathbf{A}^n \mathbf{u}^n, \quad n \in \{n_t - 1, \dots, 0\}$$

where

$$\mathbf{B}^n = (B_1^n, \dots, B_M^n)', \quad \mathbf{u}^n = (\lambda_1^n, \dots, \lambda_M^n)', \quad \mathbf{A}^n = (H_+^n(x_i) \phi(r_{im}))_{1 \leq i, m \leq M}$$

and

$$B_i^n = \sum_{m=1}^M \lambda_m^n H_-^n(x_i) \phi(r_{im}) - d^n(x_i) - d^{n-1}(x_i), \quad i = 1, \dots, M.$$

Using the terminal condition (2.10) yields

$$\mathbf{B} = \mathbf{A} \mathbf{u}^{n_t},$$

where

$$\mathbf{B} = (B_1, \dots, B_M)', \quad \mathbf{u}^{n_t} = (\lambda_1^{n_t}, \dots, \lambda_M^{n_t})', \quad \mathbf{A} = (\phi(r_{im}))_{1 \leq i, m \leq M}$$

and

$$B_i = \frac{x_i^{1-\gamma}}{1-\gamma}, \quad i = 1, \dots, M.$$

Finally, the above systems of equations can be solved by using matrix solver such as Gaussian elimination.

Sequential quadratic programming

The sequential quadratic programming (SQP) technique is applied to solve the constrained nonlinear optimization problem (2.11). We refer to NOCEDAL & WRIGHT [64] and BOMZE ET AL. [19] for a comprehensive overview of SQP methods and only briefly discuss them here. The basic idea of the SQP approach is to model the constrained nonlinear optimization problem (CNP) by a quadratic programming (QP) subproblem for a given iterate x^k . Then the QP subproblem is solved and its solution is used to construct a new iterate x^{k+1} such that the sequence $(x_k)_{k \in \mathbb{N}}$ converges to a local minimum x^* of the CNP as $k \rightarrow \infty$.

For a functional $f: \mathbb{R}^n \rightarrow \mathbb{R}$, we denote by $\nabla f(\mathbf{x})$ the gradient of f at $\mathbf{x} \in \mathbb{R}^n$ and by $Hf(\mathbf{x})$ the Hessian of f at $\mathbf{x} \in \mathbb{R}^n$, i.e.,

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)' \quad \text{and} \quad Hf(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{1 \leq i, j \leq n}.$$

For vector-valued functions $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ the symbol ∇ is used for the Jacobian of h , i.e.,

$$\nabla h(\mathbf{x}) = (\nabla h_1(\mathbf{x}), \nabla h_2(\mathbf{x}), \dots, \nabla h_m(\mathbf{x})).$$

We consider the CNP of finding a local minimizer $\mathbf{x} \in \mathbb{R}^n$ for a nonlinear objective functional $f: \mathbb{R}^n \rightarrow \mathbb{R}$ subject to a set of nonlinear equality and inequality constraints described by the functions $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$, respectively, i.e.,

$$\begin{array}{ll} \text{CNP} & \text{minimize} \quad f(\mathbf{x}) \\ & \text{over} \quad \mathbf{x} \in \mathbb{R}^n \\ & \text{subject to} \quad h(\mathbf{x}) = \mathbf{0} \\ & \quad \quad \quad g(\mathbf{x}) \leq \mathbf{0}, \end{array}$$

where $f(\mathbf{x})$, $h(\mathbf{x})$, and $g(\mathbf{x})$ are smooth functions.

Definition 2.3.2. Let $\boldsymbol{\lambda} \in \mathbb{R}^m$, $\boldsymbol{\mu} \in \mathbb{R}_*^p$, and $\mathcal{L}: \mathbb{R}^{n \times m \times p} \rightarrow \mathbb{R}$ defined by

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := f(\boldsymbol{x}) + \boldsymbol{\lambda}'h(\boldsymbol{x}) + \boldsymbol{\mu}'g(\boldsymbol{x}).$$

Then \mathcal{L} is called the Lagrangian functional of the CNP and the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are referred to as Lagrangian multipliers.

Given a current iterate $(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)$ of the CNP, the quadratic Taylor series approximation in \boldsymbol{x} for the Lagrangian functional yields

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \approx \mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) + \nabla \mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)'d(\boldsymbol{x}) + \frac{1}{2}d(\boldsymbol{x})'H\mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)d(\boldsymbol{x})$$

where $d(\boldsymbol{x}) := \boldsymbol{x} - \boldsymbol{x}^k$. The constraints are approximated with a first-order Taylor expansion yielding

$$h(\boldsymbol{x}) \approx h(\boldsymbol{x}^k) + \nabla h(\boldsymbol{x}^k)d(\boldsymbol{x}) \quad \text{and} \quad g(\boldsymbol{x}) \approx g(\boldsymbol{x}^k) + \nabla g(\boldsymbol{x}^k)d(\boldsymbol{x}).$$

Then we can write the quadratic subproblem (QSP) as

$$\begin{array}{ll} \text{QSP} & \text{minimize} \quad \nabla \mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)'d(\boldsymbol{x}) + \frac{1}{2}d(\boldsymbol{x})'H\mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)d(\boldsymbol{x}) \\ & \text{over} \quad d(\boldsymbol{x}) \in \mathbb{R}^n \\ & \text{subject to} \quad h(\boldsymbol{x}^k) + \nabla h(\boldsymbol{x}^k)d(\boldsymbol{x}) = \mathbf{0} \\ & \quad \quad \quad g(\boldsymbol{x}^k) + \nabla g(\boldsymbol{x}^k)d(\boldsymbol{x}) \leq \mathbf{0}. \end{array}$$

The QSP is equivalent to the minimization of a convex quadratic subproblem with linear constraints (c.f. BÖHME & FRANK [18, Chapter 2]), i.e.,

$$\begin{array}{ll} & \text{minimize} \quad \nabla f(\boldsymbol{x}^k)'d(\boldsymbol{x}) + \frac{1}{2}d(\boldsymbol{x})'Hf(\boldsymbol{x}^k)d(\boldsymbol{x}) \\ & \text{over} \quad d(\boldsymbol{x}) \in \mathbb{R}^n \\ & \text{subject to} \quad h(\boldsymbol{x}^k) + \nabla h(\boldsymbol{x}^k)d(\boldsymbol{x}) = \mathbf{0} \\ & \quad \quad \quad g(\boldsymbol{x}^k) + \nabla g(\boldsymbol{x}^k)d(\boldsymbol{x}) \leq \mathbf{0}. \end{array}$$

It can be solved by active set or interior-point methods. To get more information about these methods possible references are NOCEDAL & WRIGHT [64], FLETCHER [34], and WRIGHT [81]. The solution to the quadratic subproblem is regarded as search direction \boldsymbol{d}_k and the next iterate is obtained by applying the update rule

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha^k \boldsymbol{d}_k,$$

where α^k is a positive step-length.

2.3.2 Discrete-time optimization problem

The value function and optimal portfolio-proportion process of the discrete-time optimization problem are obtained by applying the following recursive algorithm.

Algorithm 2.3.3 (Backward Induction).

1. For $x > 0$ compute

$$V^N(x) = \frac{x^{1-\gamma}}{1-\gamma}.$$

2. For $x > 0$ and $k \in \{N-1, \dots, 0\}$ compute

$$V^k(x) = \sup_{(\bar{\varphi}, \bar{\eta}) \in \mathcal{K}_D^R(t_k, x)} \left\{ \mathbb{E} [V^{k+1} (e^{r\Delta} (x - \bar{\eta} + \bar{\varphi}' \cdot \mathbf{R}_{t_{k+1}}))] + \frac{\bar{\eta}^{1-\gamma}}{1-\gamma} \right\}.$$

Compute a maximizer $\boldsymbol{\nu}^k = (\varphi^k, \eta^k)$ of V^{k+1} .

3. If $k = 0$, then the value function is given by $V(t_0, x) = V^0(x)$ and the optimal portfolio-proportion process is given by $\boldsymbol{\nu} = (\boldsymbol{\nu}^0, \boldsymbol{\nu}^1, \dots, \boldsymbol{\nu}^{N-1})$. Otherwise, go back to step 2.

Theorem 2.2.2 shows that the backward induction algorithm provides the value function and optimal portfolio-proportion process. In each iteration we have to evaluate an integral and solve a constrained optimization problem. The integral is numerically evaluated by using the Gauss–Kronrod quadrature formula (see MONEGATO [62]) and the constrained nonlinear optimization problem is numerically solved by using sequential quadratic programming.

2.3.3 Financial market with a bond and a single stock

Our numerical experiments are based on the following model parameters. The financial market consists of a bond with risk-free interest rate $r = 0.1$ and a single stock with drift $\mu = 0.18$ and volatility $\sigma = 0.35$ (numerical results for other financial market parameters are not shown but are qualitatively the same). If not stated otherwise, the parameter of the power utility function is $\gamma = 0.3$ and the terminal trading time is $T = 2$ years. In the continuous-time case the dynamic risk measures are evaluated under the assumption that the portfolio-proportion process is kept constant between t and $t + \Delta$ with $\Delta = \frac{1}{24} \approx 2$ weeks. The probability level in the definition of the Value at Risk and Tail Conditional Expectation is given by $\alpha = 0.01$.

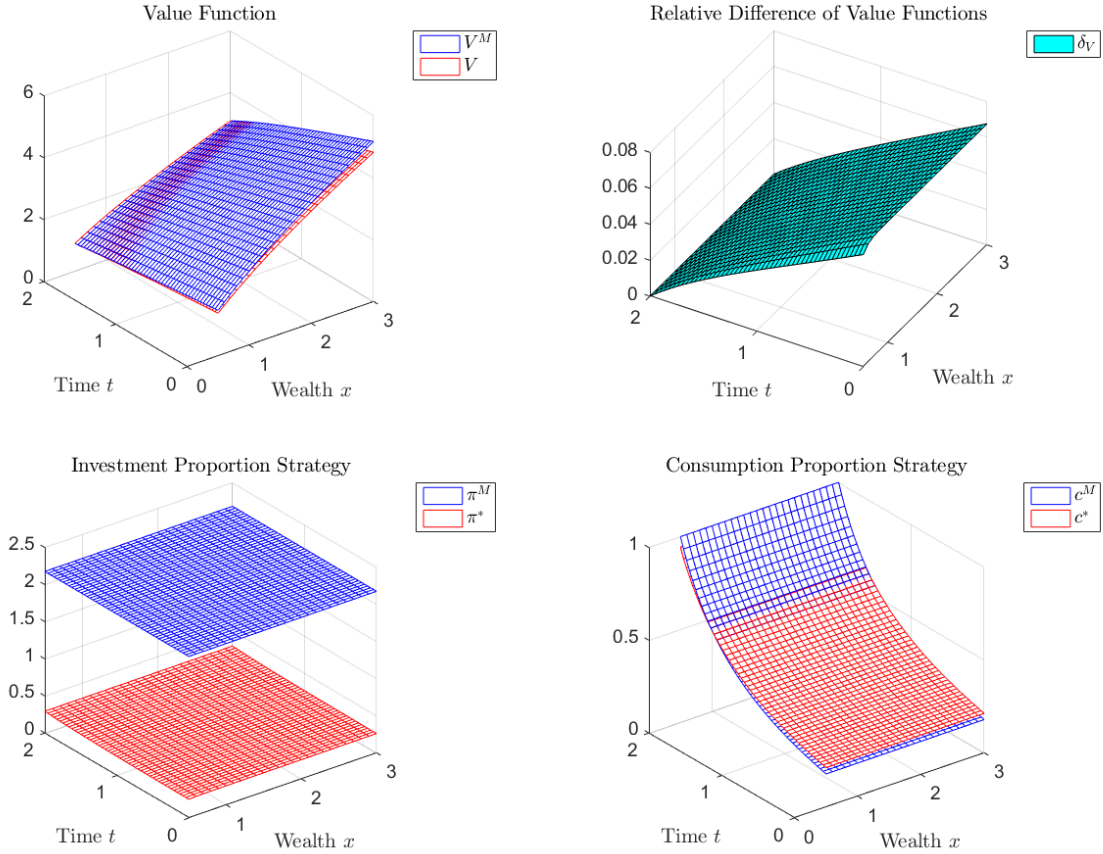


Figure 2.1: Effect of the VaR constraint on the value function and optimal portfolio-proportion strategy for the continuous-time problem. The benchmark and bound are given by $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$ and $\tilde{\varepsilon}(t, x) = 0.05x$, respectively.

Continuous-time optimization problem

We start with the continuous-time problem by numerically solving the HJB equation (2.5). Using the PI algorithm described in Section 2.3.1, we obtain an approximation of the value function and of the optimal portfolio-proportion process. Recall, in each iteration of the PI algorithm, we have to solve a linear PDE and a constrained optimization problem. Both the linear PDE and the constrained optimization problem are solved numerically. The former by using meshless methods (the multi-quadrics function is used as radial basis function with parameters $q = 1/2$, $\alpha_c = 0.1$ and $d_c = 10.3$), the latter with sequential quadratic programming methods.

Figure 2.1 shows the effect of the VaR constraint on the value function and optimal portfolio-proportion process. The benchmark for the VaR is the conditional expectation of wealth $X_{t+\Delta}^M$ obtained by an investor following the Merton portfolio-

proportion process (π_t^M, c_t^M) (see (2.6)) in $[t, t + \Delta]$ given the wealth $X_t^{(\pi, c)}$ at time t , i.e.,

$$Y_t^M := \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}] = X_t^{(\pi, c)} \exp \{ (r + \pi_t^M (\mu - r) - c_t^M) \Delta \}.$$

The bound for the VaR is given by $\tilde{\varepsilon}(t, x) = 0.05x$. This is a bound which is relative to the wealth, i.e., any loss in the interval $[t, t + \Delta]$ can be hedged with 5% of the portfolio value. For comparison, the risk of the Merton portfolio-proportion process (π_t^M, c_t^M) which is held constant in $[t, t + \Delta]$ is $\tilde{\varepsilon}(t, x) \approx 0.31x$. A first look at Figure 2.1 indicates that the value function and the relative consumption are not remarkably affected by the VaR constraint whereas the proportion of wealth invested in the risky stock is considerably reduced. The top, right-hand panel of Figure 2.1 shows the relative difference $\delta_V(t, x)$ between the value function of a Merton investor and a VaR-constrained investor defined by

$$\delta_V(t, x) = \frac{V^M(t, x) - V(t, x)}{V^M(t, x)}.$$

In order to facilitate the comparison of the value functions we express the losses of performance due to the risk constraint in monetary units and introduce the following efficiency measure. The efficiency of an investor A relative to an investor B is the initial amount of wealth that investor B would need to obtain a value function identical to that of investor A who started at time $t = 0$ with unit wealth. Figure 2.2 illustrates the efficiency of a VaR-constrained investor relative to a Merton investor in continuous time for different relative bounds of the VaR-constraint. As expected, the efficiency increases when the bound of the risk constraint becomes less restrictive. For the bound $\tilde{\varepsilon}(t, x) = 0.05x$ used in Figure 2.1 the loss of efficiency is about 9.5%. For the most restrictive case, where wealth below the benchmark is not tolerated at all, i.e., the bound is set to $\tilde{\varepsilon}(t, x) = 0$, the loss of efficiency is 12.6%.

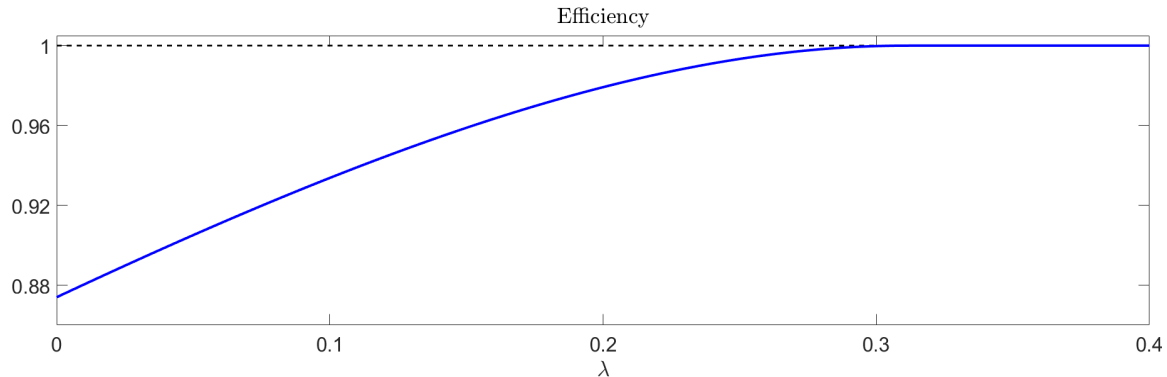


Figure 2.2: Efficiency of a VaR-constrained investor relative to a Merton investor in continuous time for different relative bounds $\tilde{\varepsilon}(t, x) = \lambda x$.

Using a relative bound leads to a constant proportion of wealth invested in the risky stock as in the unconstrained case. If we change the relative bound to an absolute one, e.g., $\tilde{\varepsilon}(t, x) \equiv 0.05$, the optimal proportion of wealth invested in the risky stock is no longer independent of the wealth level (see Figure 2.3). This results from the fact that for $x < 1$, the absolute bound is less restrictive and for $x > 1$, it is more restrictive than the relative one. For the optimal relative consumption no remarkable differences between an absolute and a relative bound are observed. The top panels

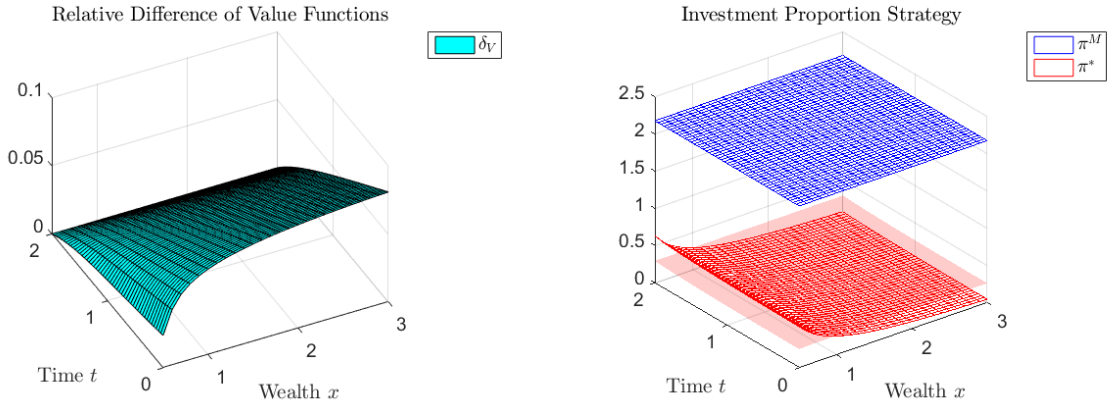


Figure 2.3: Effect of an absolute bound for the VaR constraint on the value function and optimal investment proportion strategy for the continuous-time problem. The benchmark and bound are given by $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$ and $\tilde{\varepsilon}(t, x) = 0.05$, respectively. For comparison, the optimal investment proportion strategy for $\tilde{\varepsilon}(t, x) = 0.05x$ is presented in semitransparent color.

of Figure 2.4 illustrate that a more restrictive bound leads to a smaller investment in the risky stock and a smaller value function. Another interesting observation can be made from the bottom panels of Figure 2.4. The value function at time $t = 0$ is plotted against the wealth level for different terminal trading times T . The left panel shows the value function of a Merton investor (solid line), of a VaR-constrained investor (dashed line) and a TCE-constrained investor (asterisk). In the right panel, the value function of a Merton investor (solid line) and of a EL-constrained investor (dashed line) is plotted. We observe that the value function is not noticeably affected by the choice of different risk constraints and that there are almost no differences between the VaR constraint and the TCE constraint. This confirms similar results by CUOCO ET AL. [24]. Even though a static VaR constraint has been found to induce an increased probability of extreme losses and an increased allocation to risky assets in some states (see BASAK & SHAPIRO [5]), these shortcomings vanish if a VaR constraint is imposed dynamically (see [24]).

For longer trading horizons the effects of a risk constraint on the value function become more noticeable. We obtain comparable results if we change the benchmark

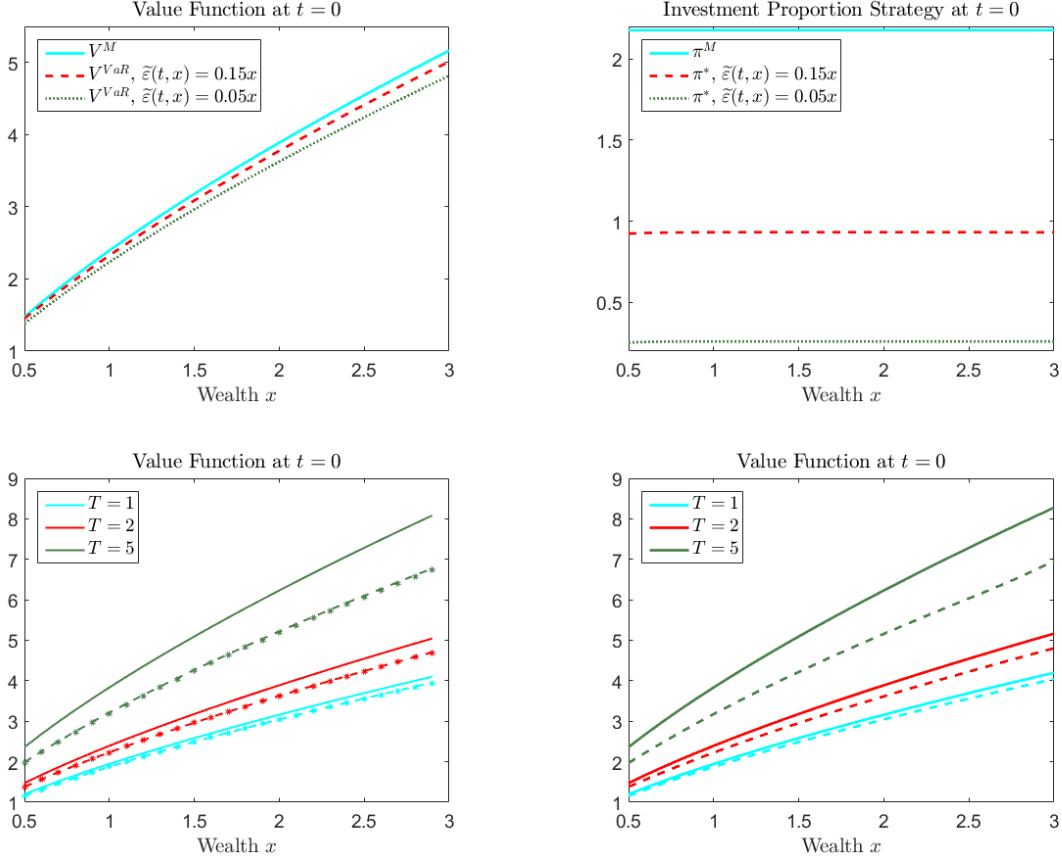


Figure 2.4: **Top panels:** Effect of different relative bounds ($\tilde{\varepsilon}(t, x) = 0.15x, 0.05x$) for the VaR constraint on the value function and optimal investment proportion strategy in the continuous-time case. The benchmark and terminal trading time are given by $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$ and $T = 2$, respectively.

Bottom panels: Effect of different terminal trading times ($T = 1, 2, 5$) and different risk constraints (VaR, TCE, EL) on the value function in the continuous-time case. The benchmark is given by $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$.

Bottom left panel: Value function of the Merton investor (solid line), the VaR-constrained investor (dashed line) and the TCE-constrained investor (asterisk) for $\tilde{\varepsilon}(t, x) = 0.05x$.

Bottom right panel: Value function of the Merton investor (solid line) and the EL-constrained investor (dashed line) for $\tilde{\varepsilon}(t, x) = 0.01x$.

to the wealth an investor will obtain at time $t + \Delta$ starting with $X_t^{(\pi, c)}$ at time t and only investing in the bond (and not in the stock) and consuming the wealth with the rate c_t^M in $[t, t + \Delta]$, i.e. $Y_t = X_t^{(\pi, c)} \cdot e^{(r - c_t^M)\Delta}$.

Remark 2.3.4. We also performed numerical experiments for the case $\gamma > 1$ instead of $\gamma = 0.3$ which led to smaller differences between the value function of a Merton

investor and a risk-constrained investor. This results from the fact that for a larger γ an investor is more risk-averse even without an imposed risk constraint.

Discrete-time optimization problem

We now consider the discrete-time problem and apply the backward induction algorithm described in Section 2.3.2. The expectation in (2.8) can be written as

$$\begin{aligned} & \mathbb{E} \left[(1 + \bar{\beta} \cdot R_{t_{n+1}})^{1-\gamma} \right] \\ &= \int_0^\infty \left[(1 + \bar{\beta}(u-1))^{1-\gamma} \cdot \frac{1}{\sqrt{2\pi\Delta}\sigma u} \exp \left\{ -\frac{(\ln u - (\mu - r - \frac{\sigma^2}{2})\Delta)^2}{2\sigma^2\Delta} \right\} \right] du, \end{aligned}$$

where we used that $R_{t_{n+1}} = \tilde{R}_{t_{n+1}}/e^{r\Delta} - 1$ and $\tilde{R}_{t_{n+1}}$ is log-normally distributed. Figure 2.5 shows the effect of the VaR constraint on the value function and optimal

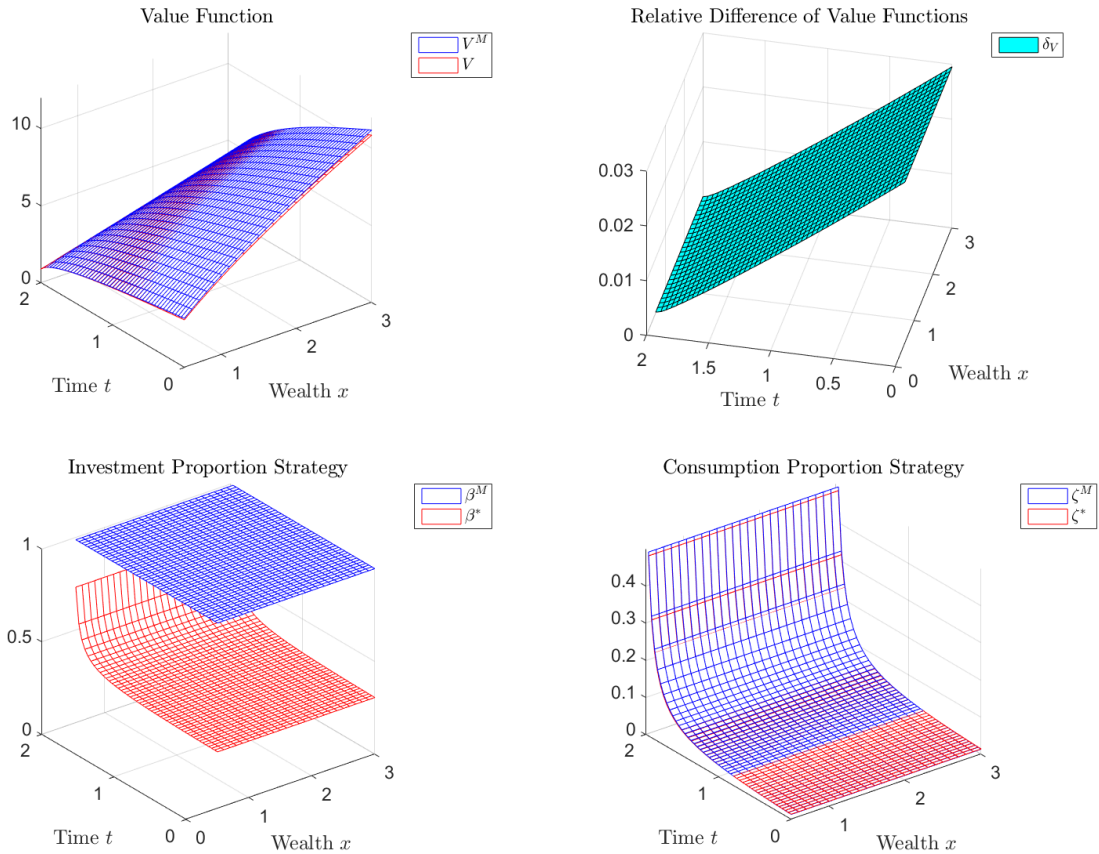


Figure 2.5: Effect of the VaR constraint on the value function and optimal portfolio-proportion strategy in the discrete-time case. The benchmark and bound are given by $Y_{t_n}^M = \mathbb{E}[X_{t_{n+1}}^M | X_{t_n}^{(\varphi, \eta)}]$ and $\tilde{\epsilon}(t_n, x) = 0.05x$, respectively.

strategy. The benchmark for the VaR is the conditional expectation of wealth $X_{t_{n+1}}^M$ obtained by an investor following the (discrete-time) Merton strategy $(\varphi_{t_n}^M, \eta_{t_n}^M)$ given the wealth $X_{t_n}^{(\varphi, \eta)}$ at time t_n , i.e.,

$$Y_{t_n}^M = \mathbb{E}[X_{t_{n+1}}^M | X_{t_n}^{(\varphi, \eta)}] = e^{r\Delta} \left(X_{t_n}^{(\varphi, \eta)} - \eta_{t_n}^M - \varphi_{t_n}^M \right) + e^{\mu\Delta} \varphi_{t_n}^M.$$

The bound for the VaR constraint was given by $\tilde{\epsilon}(t_n, x) = 0.05x$. In comparison, the risk of the Merton strategy $(\varphi_{t_n}^M, \eta_{t_n}^M)$ is $\tilde{\epsilon}(t, x) \approx 0.16x$. It can be observed that the value function is not remarkably affected by the VaR constraint. Moreover, the effects on the value function are even less notable than in the continuous-time case. This results from the fact that even without an imposed risk constraint short-selling the stock or bond is not allowed, thus the proportion invested in the stock is always in $[0, 1]$. Note that in the above example for the continuous-time problem we have for the Merton fraction $\pi^M \approx 2.18$, i.e., it exceeds one. Furthermore, it can be observed from Figure 2.5 that the fraction of wealth invested in the risky stock is considerably reduced when the VaR constraint is imposed, whereas the differences in the consumption rate between a VaR-constraint investor and a Merton investor are hard to be visually distinguished. If we use an absolute bound $\tilde{\epsilon}(t, x) \equiv 0.05$ for the VaR constraint instead of a relative bound, the optimal investment strategy is no longer a constant proportion of wealth and the results are similar to Figure 2.3 for the continuous-time case. Numerical results for varying the terminal trading time or the risk measure are not shown here, but they are comparable to the continuous-time case (see Figure 2.4).

As in the continuous-time case we now express the losses of performance of a risk-constrained investor relative to the performance of a Merton investor in monetary units using the efficiency measure. The left panel of Figure 2.6 illustrates the efficiency of a VaR-constrained investor relative to a Merton investor in discrete time for different relative bounds of the VaR-constraint. As expected and already observed in Figure 2.2, the efficiency increases when the bound of the risk constraint becomes less restrictive. The loss of efficiency is at most 7.2% which is attained for the most restrictive bound, i.e., $\tilde{\epsilon}(t, x) = 0$. For the bound $\tilde{\epsilon}(t, x) = 0.05x$ used in Figure 2.5 the loss of efficiency is about 4.2%.

We finish this section with a numerical comparison of the continuous-time and discrete-time case. The availability of solutions to the risk-constrained portfolio problem both for discrete and continuous time allows us to quantify the losses of portfolio performance resulting from time discretization, i.e., from restricting to finite trading frequencies. First observe that the loss of efficiency resulting from imposing a risk-constraint is higher in continuous time than in discrete time (see Figure 2.2 and left panel of Figure 2.6). This arises from the fact that short-selling is allowed in continuous time leading to $\pi^M \approx 2.18 > 1$ whereas it is not allowed in discrete time leading to $\beta^M = 1$. For a fair comparison we do not allow for short-selling in both cases. Since the parameters of the financial market shall remain the same, we change

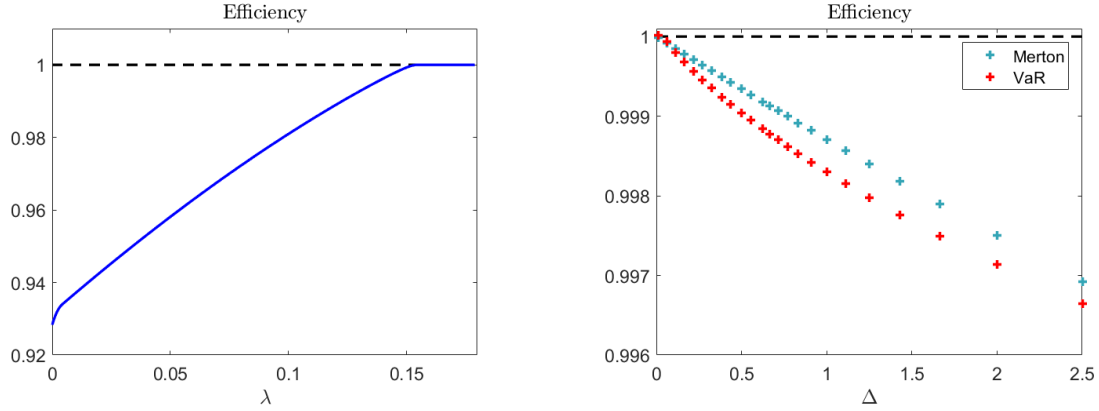


Figure 2.6: Efficiency for different bounds $\tilde{\epsilon}(t, x) = \lambda x$ (left panel) and for different risk measurement horizons Δ (right panel). The benchmark is given by Y_t^M .

Left panel: Efficiency of a VaR-constrained investor relative to a Merton investor in discrete time for $\gamma = 0.3$ and $\Delta = 1/24$.

Right panel: Efficiency of a discrete-time relative to a continuous-time Merton investor and efficiency of a discrete-time relative to a continuous-time VaR-constrained investor for $\gamma = 0.9$ and $\tilde{\epsilon}(t, x) = 0.05x$.

the investors' preferences represented by the utility functions by setting $\gamma = 0.9$. This leads to $\pi^M, \beta^M \in [0, 1]$, i.e., there is no short-selling in both cases. Furthermore, the expected utility from consumption in the time interval $[t_n, t_{n+1})$, $n = 0, \dots, N - 1$, is given by $\mathbb{E}[\int_{t_n}^{t_{n+1}} U_1(C_s) ds]$ in the continuous-time case. However, in the discrete-time case we have $\mathbb{E}[U_1(\eta_{t_n})]$, which approximately corresponds to $\mathbb{E}[U_1(\int_{t_n}^{t_{n+1}} C_s ds)]$. The related maximization problem

$$\mathbb{E}_{t_0, x_0} \left[U_1 \left(\int_0^T C_t dt \right) + U_2(X_T^u) \right]$$

is not only different from the economic interpretation, but also from the mathematical point of view (see GRANDITS ET AL. [38]). Since we want to quantify the losses of portfolio performance resulting solely from restricting to finite trading frequencies, we set $U_1(x) \equiv 0$. Again, we use the efficiency to compare the portfolio performance of a discrete-time investor relative to a continuous-time investor with and without imposing a risk constraint. For this experiment, we compute the VaR risk measure for the continuous-time investor under the assumption that the number of shares remains constant between t and $t + \Delta$. In ROGERS [74] the continuous-time investment problem is compared to the discrete-time investment problem when no risk constraint is imposed. The author shows that there is only a small difference between the continuous-time and discrete-time problem if the additional restriction is imposed that the continuous-time investor is not allowed to sell short the stock or the bond.

BÄUERLE ET AL. [9] show that this remains true for models where the drift is modeled by a random variable which is not directly observable and has to be estimated from observed stock prices. However, the authors found that if the short-selling restriction for the continuous-time investor is omitted, then the discrete-time investor cannot in general do as well as the continuous-time one and a discretization gap remains. The right panel of Figure 2.6 shows the efficiency of a discrete-time relative to a continuous-time Merton investor (cyan) and the efficiency of a discrete-time relative to a continuous-time VaR-constrained investor (red). We observe that the values of the efficiency are very close to unity; even for a quite large Δ in the range of 2.5 years, the loss of efficiency is at most 0.35% in both cases. This is in line with the results obtained in [74] for the Merton investor. In addition to the results in [74] our numerical results for the risk-constrained case indicate that the losses due to time discretization are of comparable (small) magnitude.

2.3.4 Financial market with a bond and two stocks

In this section we consider a continuous-time financial market consisting of one bond and two stocks. We start with investigating the impact of the correlation coefficient on the optimal investment proportion strategy of the Merton investor. The prices of the stocks are modeled by the SDEs

$$\begin{aligned} dS_t^i &= S_t^i \left(\mu_i dt + \sum_{j=1}^2 \sigma_{ij} dW_t^j \right), \quad i = 1, 2 \\ S_0^i &= s^i, \end{aligned}$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

are the drift and volatility, respectively. We assume that the volatility matrix is positive definite and we denote the covariance matrix of the log returns of the stocks by $\boldsymbol{\Sigma}$ which is given by

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11}^2 + \sigma_{12}^2 & \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} \\ \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} & \sigma_{21}^2 + \sigma_{22}^2 \end{pmatrix}.$$

The volatility matrix $\boldsymbol{\sigma}$ equals the square root of the covariance matrix $\boldsymbol{\Sigma}$ and a typical choice for $\boldsymbol{\sigma}$ is the Cholesky decomposition matrix of $\boldsymbol{\Sigma}$, i.e., $\bar{\boldsymbol{\sigma}}\bar{\boldsymbol{\sigma}}' = \boldsymbol{\Sigma}$, where $\bar{\boldsymbol{\sigma}}$ is lower triangular. Hence, we define the volatility matrix as

$$\bar{\boldsymbol{\sigma}} = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}, \quad (2.16)$$

where $\sigma_1 := \sqrt{(\sigma_{11})^2 + (\sigma_{12})^2}$, $\sigma_2 := \sqrt{(\sigma_{22})^2 + (\sigma_{21})^2}$ and $\rho \in (-1, 1)$ denotes the correlation coefficient of the stocks. We denote the optimal investment proportion strategy of the Merton investor by $\boldsymbol{\pi}_\rho^M$ and by Equation (2.6) we have

$$\boldsymbol{\pi}_\rho^M = \begin{pmatrix} \pi_\rho^{M,1} \\ \pi_\rho^{M,2} \end{pmatrix} = \frac{1}{(1 - \rho^2)\gamma} \begin{pmatrix} \frac{\mu_1 - r}{(\sigma_1)^2} - \rho \frac{(\mu_2 - r)}{\sigma_1 \sigma_2} \\ \frac{\mu_2 - r}{(\sigma_2)^2} - \rho \frac{(\mu_1 - r)}{\sigma_1 \sigma_2} \end{pmatrix}. \quad (2.17)$$

For $\rho = 0$ we obtain

$$\boldsymbol{\pi}_0^M = \begin{pmatrix} \pi_0^{M,1} \\ \pi_0^{M,2} \end{pmatrix} = \begin{pmatrix} \pi_{d=1}^{M,1} \\ \pi_{d=1}^{M,2} \end{pmatrix},$$

where

$$\pi_{d=1}^{M,i} = \frac{\mu_i - r}{\gamma \sigma_i^2}.$$

It holds that $\pi_{d=1}^{M,i}$ equals the optimal investment proportion strategy of the Merton investor given a financial market consisting of a bond and stock i . Thus, if the two stocks are uncorrelated, the optimal investment proportion in stock i equals the proportion invested in stock i given a financial market only consisting of a bond and stock i , for $i = 1, 2$.

For the limiting cases $\rho = +1$, i.e., the stocks have a perfect positive correlation, and $\rho = -1$, i.e., the stocks have a perfect negative correlation, the volatility matrix is no longer positive definite and the market model as well as the optimal control problem degenerate. However, it is interesting to study the limiting behavior for $\rho \rightarrow +1$ and $\rho \rightarrow -1$. For $\rho \in (-1, +1)$ we can rewrite Equation (2.17) as

$$\boldsymbol{\pi}_\rho^M = \begin{pmatrix} \pi_\rho^{M,1} \\ \pi_\rho^{M,2} \end{pmatrix} = \frac{1}{(1 - \rho)(1 + \rho)\gamma} \begin{pmatrix} \frac{\mu_1 - r}{\sigma_1^2} \left(1 - \rho \frac{(\mu_2 - r)\sigma_1}{(\mu_1 - r)\sigma_2}\right) \\ \frac{\mu_2 - r}{\sigma_2^2} \left(1 - \rho \frac{(\mu_1 - r)\sigma_2}{(\mu_2 - r)\sigma_1}\right) \end{pmatrix}.$$

Assuming it holds that

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}, \quad (2.18)$$

we obtain

$$\boldsymbol{\pi}_\rho^M = \begin{pmatrix} \pi_\rho^{M,1} \\ \pi_\rho^{M,2} \end{pmatrix} = \frac{1}{1 + \rho} \begin{pmatrix} \frac{\mu_1 - r}{\gamma \sigma_1^2} \\ \frac{\mu_2 - r}{\gamma \sigma_2^2} \end{pmatrix} = \frac{1}{1 + \rho} \begin{pmatrix} \pi_{d=1}^{M,1} \\ \pi_{d=1}^{M,2} \end{pmatrix}. \quad (2.19)$$

The term $(\mu_i - r)/\sigma_i$ in Equation 2.18 is called the Sharpe ratio of stock i and can be interpreted as a performance measure examining how well the return of stock i compensates the investor for the risk taken. In Equation 2.18 we assume that the two stocks have the same Sharpe ratio.

If ρ tends to 1, we obtain from Equation (2.19)

$$\boldsymbol{\pi}_1^M = \begin{pmatrix} \pi_1^{M,1} \\ \pi_1^{M,2} \end{pmatrix} \rightarrow \frac{1}{2} \begin{pmatrix} \pi_{d=1}^{M,1} \\ \pi_{d=1}^{M,2} \end{pmatrix},$$

whereas if ρ tends to -1 , we obtain

$$\pi_{-1}^{M,i} \rightarrow \text{sgn}(\mu_1 - r)\infty \quad \text{for } i = 1, 2.$$

Thus, if the correlation of the two stocks goes to 1 (i.e., a perfect positive correlation), the optimal investment proportion in stock i tends to half of the proportion invested in stock i given the financial market only consists of a bond and stock i . If the correlation of the two stocks goes to -1 (i.e., a perfect negative correlation), the optimal investment proportion in each stock tends to infinity. The latter can be explained as follows. By setting $\rho = -1$ in Equation (2.16), we obtain

$$\bar{\boldsymbol{\sigma}} = \begin{pmatrix} \sigma_1 & 0 \\ -\sigma_2 & 0 \end{pmatrix},$$

and from Equation (2.1) follows that the associated wealth equation satisfies

$$dX_t = X_t (r + \tilde{\pi}^1(\mu_1 - r) + \tilde{\pi}^2(\mu_2 - r)) dt + X_t(\tilde{\pi}^1\sigma_1 - \tilde{\pi}^2\sigma_2)dW_t^1,$$

where we have set $c \equiv 0$ for simplicity and $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}^1, \tilde{\pi}^2)$ denotes the investment strategy. By assuming that

$$\frac{\tilde{\pi}^1}{\tilde{\pi}^2} = \frac{\sigma_2}{\sigma_1},$$

we obtain that the wealth is governed by the following deterministic ordinary differential equation (ODE)

$$dX_t = X_t (r + \tilde{\pi}^1(\mu_1 - r) + \tilde{\pi}^2(\mu_2 - r)) dt. \quad (2.20)$$

A solution to the ODE (2.20) starting from x_0 at 0 is given by

$$X_t = x_0 \exp [(r + \tilde{\pi}^1(\mu_1 - r) + \tilde{\pi}^2(\mu_2 - r)) t].$$

By sending $\tilde{\pi}^1 \rightarrow \text{sgn}(\mu_1 - r)\infty$ and $\tilde{\pi}^2 \rightarrow \text{sgn}(\mu_2 - r)\infty$, we have $X_t \rightarrow \infty$ and also $U(X_t) \rightarrow \infty$ with U a utility function. Hence, for a perfect negative correlation of the

stocks the risk (i.e., the stochastic change) can be perfectly eliminated and unlimited profit is possible. Thus, if the correlation of the two stocks goes to -1 , the optimal investment proportion in each stock tends to infinity.

In the following we present some numerical experiments similar to those in PFÜTZNER [66]. First we investigate the impact of the stocks' correlation coefficient on the value function and the portfolio-proportion strategy of a Merton investor and a risk constrained investor, respectively. Then, we examine the impact of various benchmarks, and finally we study the impact of various bounds.

The parameters of the securities are $r = 0.1$ for the bond, $\mu^1 = 0.18$, $\sigma^1 = 0.35$ and $\mu^2 = 0.20$, $\sigma^2 = 0.45$ for the first stock and the second stock, respectively. Note that

$$(\mu^1 - r)/\sigma^1 \approx (\mu^2 - r)/\sigma^2.$$

The dynamic risk measures are evaluated under the assumption that the portfolio-proportion process is kept constant between t and $t + \Delta$ with $\Delta = \frac{1}{24} \approx 2$ weeks. The probability level in the definition of the Value at Risk and Tail Conditional Expectation is given by $\alpha = 0.01$. The parameter of the power utility function is $\gamma = 0.3$ and the terminal trading time is $T = 2$ years.

Variation of the correlation coefficient

In this section the effect of the correlation coefficient of the stocks is examined by numerical examples presented in the following figures. The VaR is used as a risk constraint with benchmark and bound given by

$$Y_t^M := \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi,c)}] \quad \text{and} \quad \tilde{\varepsilon}(t, x) = 0.05x.$$

Figure 2.7 shows the effect of the VaR constraint on the value function and optimal investment proportion strategy for different values of ρ . The left panels present the value function of a Merton investor and a VaR-constrained investor, whereas the right panels show the associated optimal investment proportion strategies. It results from the considerations above that the risk can be perfectly eliminated if the stocks have a perfect negative correlation and $\pi_{-1}^{M,1}/\pi_{-1}^{M,2} = \sigma_2/\sigma_1$ holds. Hence, by sending the investment proportion in each stock to infinity unlimited profit is possible. This is illustrated by the top panel of Figure 2.7. If the stocks have a strong negative correlation the Merton investor puts large proportions into both stocks such that

$$\frac{\pi_{-0.999}^{M,1}}{\pi_{-0.999}^{M,2}} = 1.2857 = \frac{\sigma_2}{\sigma_1}.$$

Thus, the associated value function of the Merton investor also has high values. The VaR-constrained investor chooses the same investment proportion strategy as the

Merton investor because the risk constraint is not binding due to the eliminated stochastic change. Hence the value function of a VaR-constraint investor equals the value function of a Merton investor. If the correlation coefficient increases, the Merton investor and the VaR-constrained investor put less large proportions into both stocks. Furthermore, the VaR-constrained investor invests smaller proportions into both stocks compared to the Merton investor in order to satisfy the risk constraint. If the stocks are uncorrelated we have (see also Figure 2.1)

$$\boldsymbol{\pi}_0^M = \begin{pmatrix} \pi_0^{M,1} \\ \pi_0^{M,2} \end{pmatrix} = \begin{pmatrix} \pi_{d=1}^{M,1} \\ \pi_{d=1}^{M,2} \end{pmatrix}$$

and if the stocks have a strong positive correlation we obtain

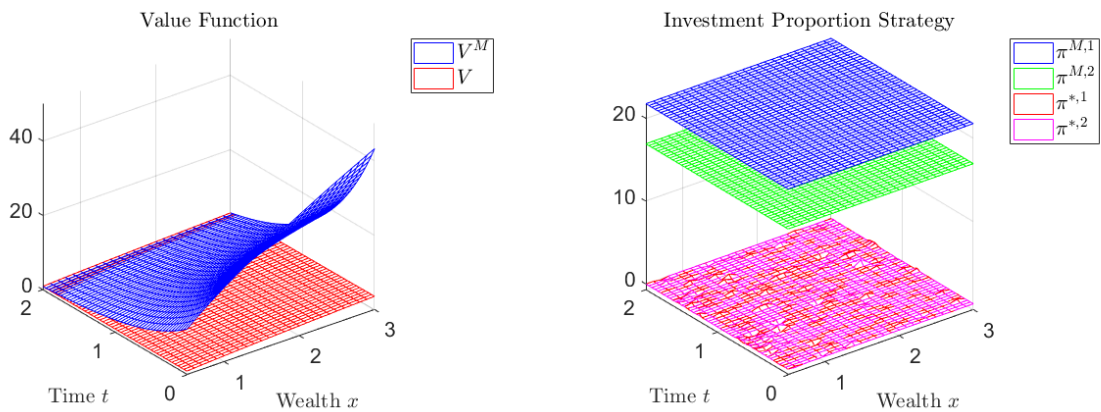
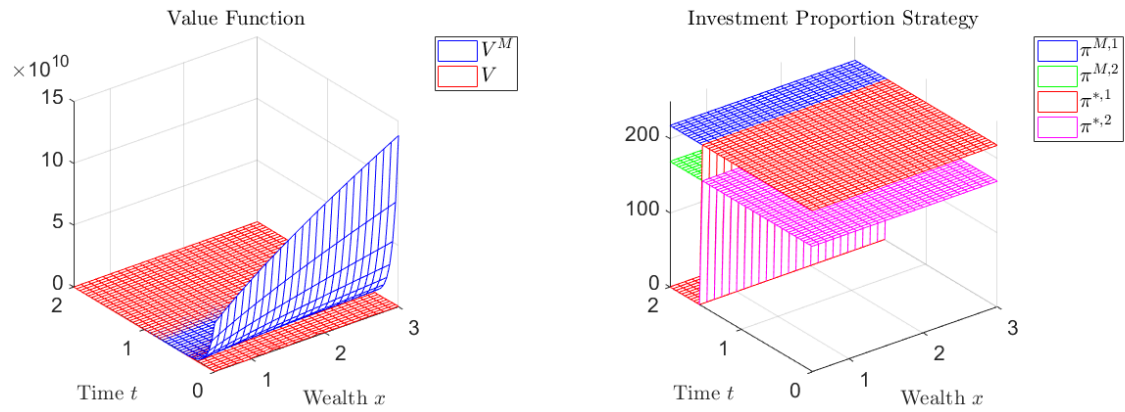
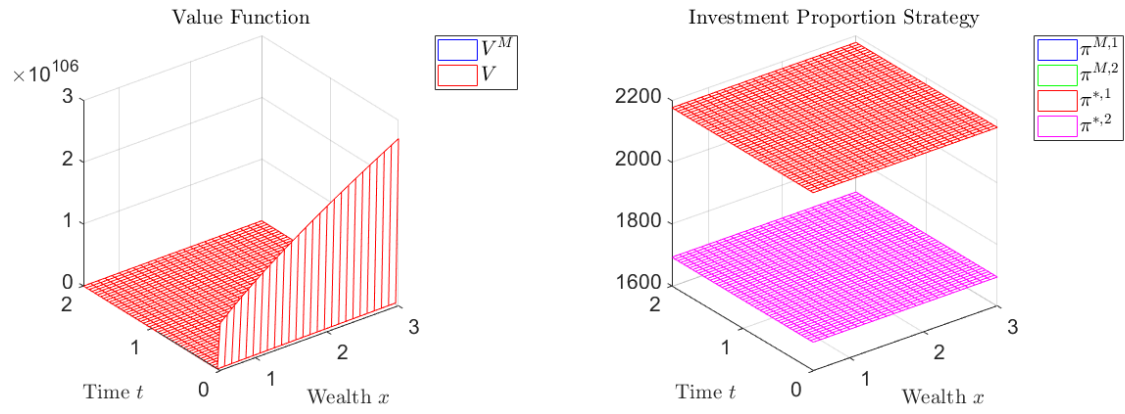
$$\boldsymbol{\pi}_{0.999}^M = \begin{pmatrix} \pi_{0.999}^{M,1} \\ \pi_{0.999}^{M,2} \end{pmatrix} \approx \frac{1}{2} \begin{pmatrix} \pi_{d=1}^{M,1} \\ \pi_{d=1}^{M,2} \end{pmatrix}.$$

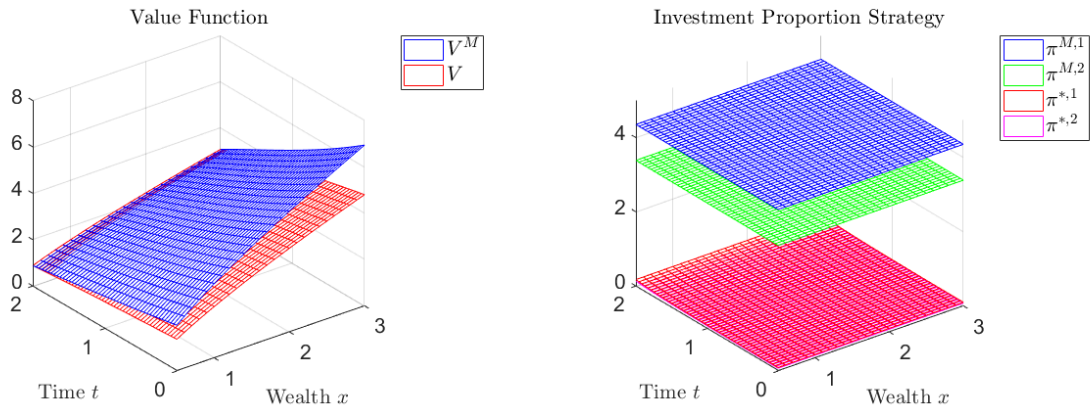
The numerical results also show that there is no analogue for a risk-constrained investor, i.e.,

$$\boldsymbol{\pi}_0^* = \begin{pmatrix} \pi_0^{*,1} \\ \pi_0^{*,2} \end{pmatrix} \neq \begin{pmatrix} \pi_{d=1}^{*,1} \\ \pi_{d=1}^{*,2} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\pi}_{0.999}^* = \begin{pmatrix} \pi_{0.999}^{*,1} \\ \pi_{0.999}^{*,2} \end{pmatrix} \neq \frac{1}{2} \begin{pmatrix} \pi_{d=1}^{*,1} \\ \pi_{d=1}^{*,2} \end{pmatrix},$$

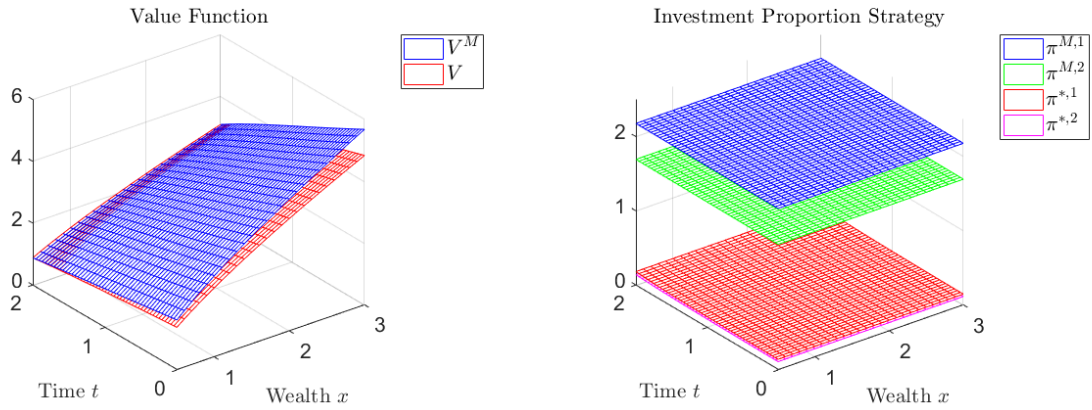
where $\pi_{d=1}^{*,i}$, $i = 1, 2$, equals the optimal investment proportion strategy of the VaR-constrained investor given a financial market only consisting of a bond and stock i . The effect of the correlation coefficient on the value function and optimal investment proportion strategy at $t = 0$ and $x = 1$ is presented in more detail in the top panel of Figure 2.8. The left panel presents the value function of a Merton investor and a VaR-constrained investor, whereas the right panel shows the associated optimal investment proportion strategies. As already shown, if the correlation coefficient tends to -1 the proportion invested into both stocks by a Merton investor tends to infinity, as well as the associated value function. The optimal investment proportion strategy and associated value function of a VaR-constrained investor also tend to infinity if ρ tends to -1 , since the risk can be perfectly eliminated and unlimited profit is possible. However, if ρ is about -0.9 we already observed from the above figures that the Merton investor puts a large proportion into both stocks, whereas the VaR-constrained investor only puts a small proportion in both stocks to satisfy the risk constraint. This leads to large differences between the value functions. If ρ increases to $+1$ the Merton investor puts smaller proportions into both stocks yielding smaller differences between the value functions. The corresponding efficiency of a VaR-constrained investor relative to a Merton investor for different values of ρ is presented in the bottom panel of Figure 2.8. If the stocks have a perfect negative correlation, the efficiency is one because the value function of the VaR-constrained

investor equals the value function of the Merton investor. However, if the stocks have a strong negative correlation the efficiency is close to zero. The efficiency considerably increases, if the correlation coefficient increases yielding an efficiency of 0.9 if the stocks have a strong positive correlation.

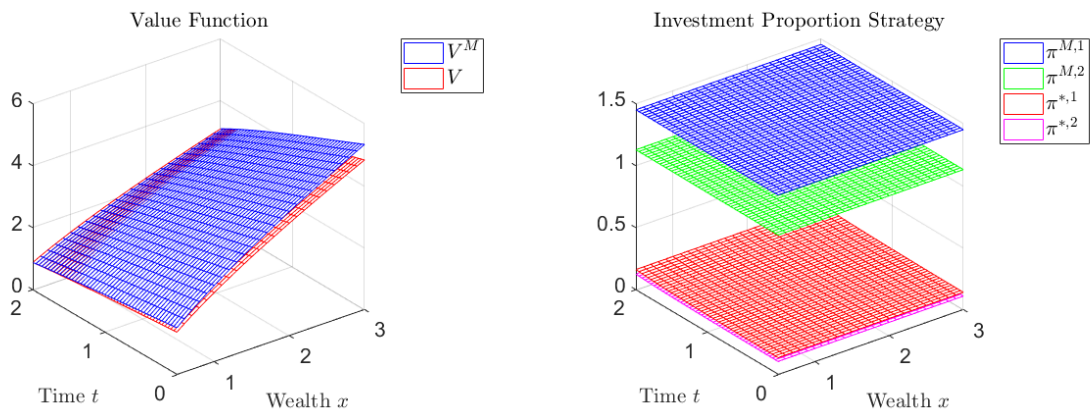




(d) $\rho = -0.5$



(e) $\rho = 0$



(f) $\rho = 0.5$

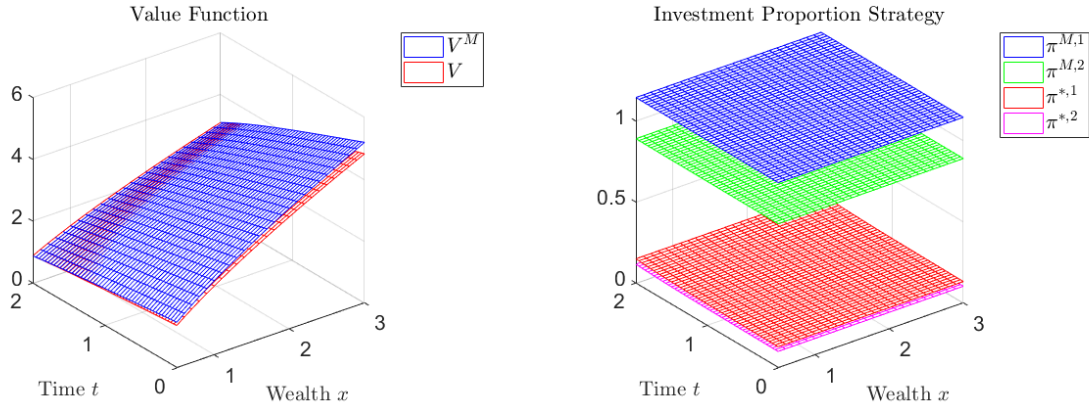
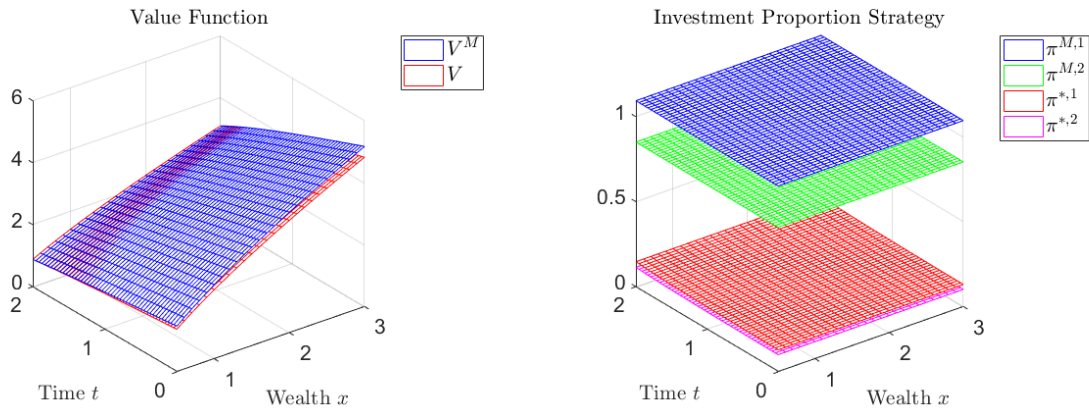
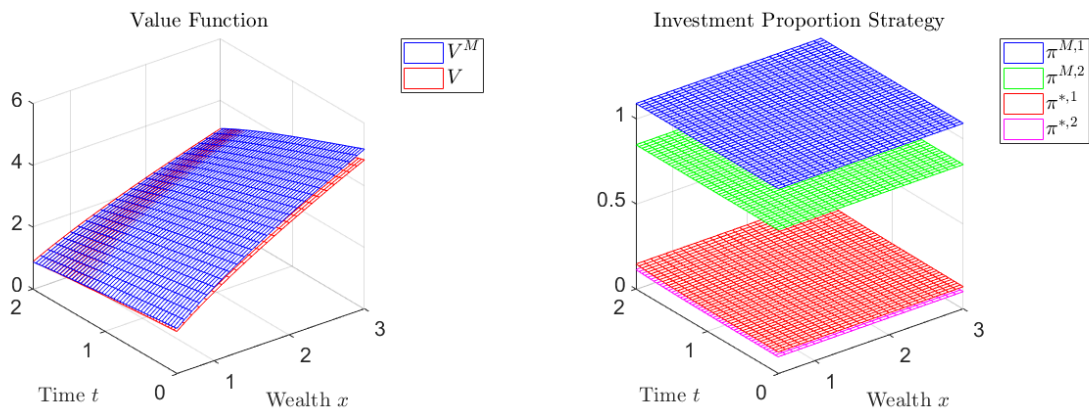
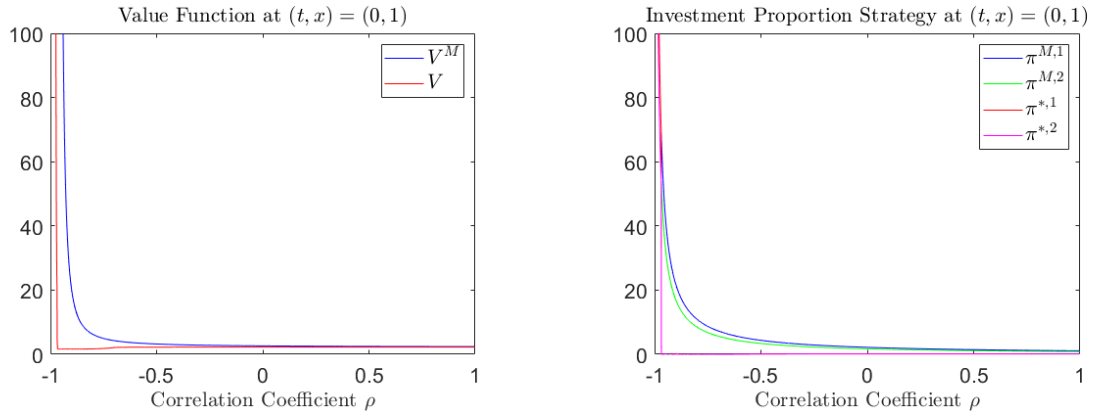
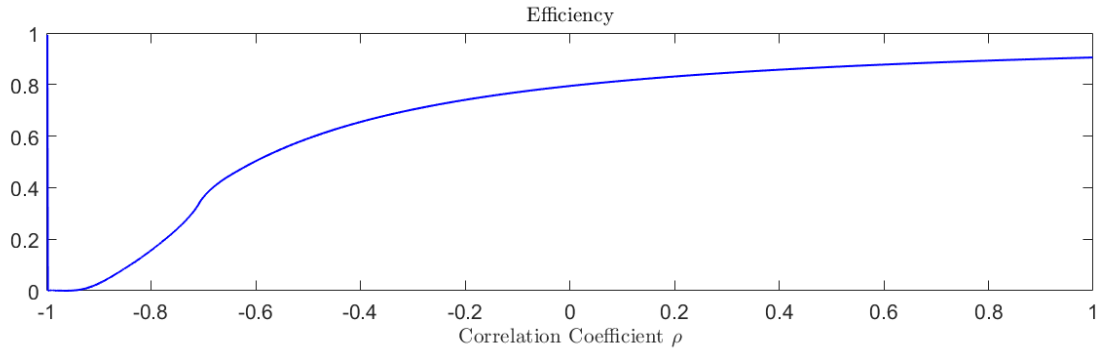
(g) $\rho = 0.9$ (h) $\rho = 0.99$ (i) $\rho = 0.999$

Figure 2.7: Effect of the VaR constraint on the value function and optimal investment proportion strategy for different values of ρ . The benchmark and bound are given by $Y_t^M := \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi,c)}]$ and $\tilde{\varepsilon}(t, x) = 0.05x$, respectively.



(a) Value function and optimal investment proportion strategy at $t = 0$ and $x = 1$ for different values of ρ .



(b) Efficiency of a VaR-constrained investor relative to a Merton investor for different values of ρ .

Figure 2.8: Effect of the VaR constraint on the value function, optimal investment proportion strategy, and efficiency for different values of ρ . The benchmark and bound are given by $Y_t^M := \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$ and $\tilde{\varepsilon}(t, x) = 0.05x$, respectively.

Variation of the benchmark

In the following numerical examples we investigate the effect of the benchmark of a risk constraint on the value function and optimal investment proportion strategy. The TCE is used as a risk constraint with the relative bound $\tilde{\varepsilon}(t, x) = 0.05x$. The correlation coefficient of the stocks is given by $\rho = 0.5$.

Figure 2.9 shows the effect of the TCE constraint on the value function and optimal investment proportion strategy for different benchmarks. The left panels present the value function of a Merton investor and TCE-constrained investor, whereas the right panels show the associated optimal investment proportion strategies. The considered benchmarks are $Y_t^M := \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$, $Y_t^B := \mathbb{E}[X_{t+\Delta}^{(0, c^M)} | X_t^{(\pi, c)}]$, and $Y_t^{0.95} := 0.95X_t^{(\pi, c)}$. The investor who is constrained to the benchmark $Y_t^{0.95}$ puts the largest proportion into both stocks, whereas the investor who is constrained to the benchmark Y_t^M puts the smallest proportion into both stocks. The differences between the value function of a Merton investor and a TCE-constrained investor are larger for the benchmark Y_t^M than for the benchmarks Y_t^B and $Y_t^{0.95}$. The effect of the benchmark on the value function and optimal investment proportion strategy is presented in more detail in Figure 2.10. The top panel shows the value function for the Merton investor and the value function of a TCE-constrained investor with the benchmarks Y_t^M , Y_t^B and $Y_t^{0.95}$, respectively. The bottom panels show the associated optimal investment proportion strategies. We observe that the benchmarks affect the value function. However the difference between the value functions due to the different benchmarks is small.

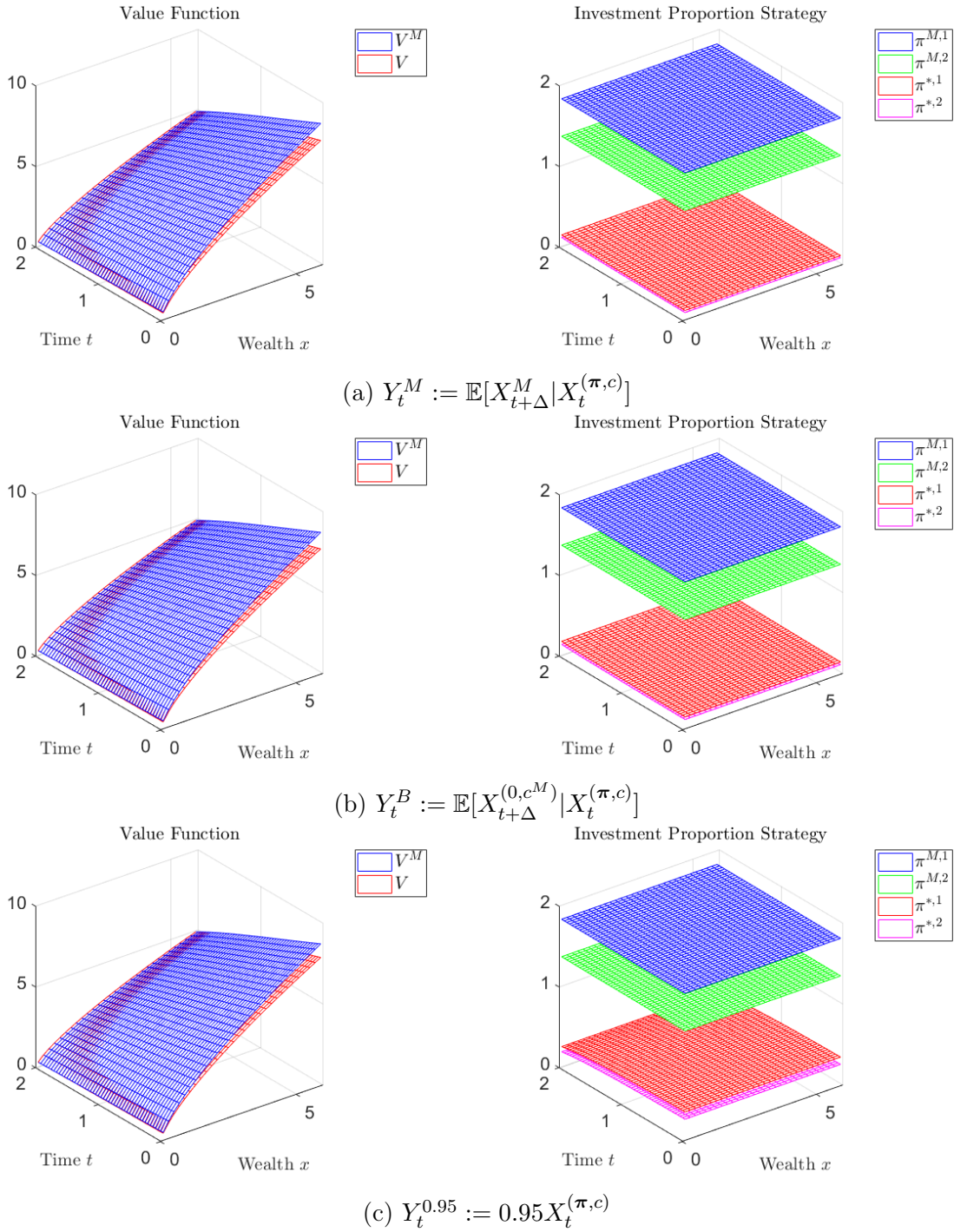


Figure 2.9: Effect of the TCE constraint on the value function and optimal investment proportion strategy for different benchmarks Y . The bound is $\tilde{\varepsilon}(t, x) = 0.05x$.

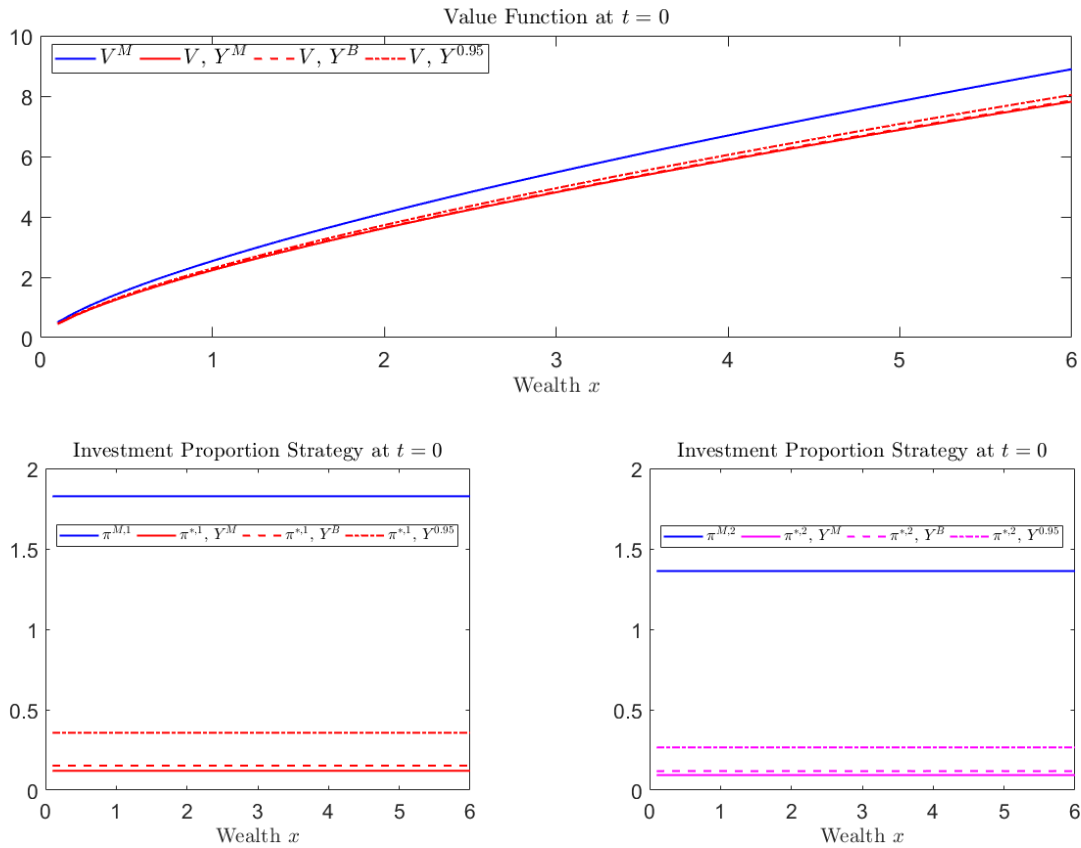


Figure 2.10: Effect of the TCE constraint on the value function at $t = 0$ and optimal investment proportion strategy at $t = 0$ for different benchmarks. The bound is $\tilde{\varepsilon}(t, x) = 0.05x$.

Variation of the bound

In the following numerical examples we investigate the effect of the bound of a risk constraint on the value function and optimal investment proportion strategy. The correlation coefficient is given by $\rho = 0.5$ and the EL is used as a risk constraint with the benchmark $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$. Figure 2.11 shows the effect of the EL constraint on the value function and optimal investment proportion strategy for different relative bounds. The left panels present the value function of a Merton investor and a EL-constrained investor, whereas the right panels show the associated optimal investment proportion strategies. It is illustrated that a more restrictive bound is associated with smaller proportions invested into the stocks. Figure 2.12 displays the effect of the EL constraint on the value function and optimal investment proportion strategy for different absolute bounds. Similar to the model with a single stock, we observe that for an absolute bound the optimal investment proportion is no longer independent of the wealth level x .

The effect of the bound on the value function and optimal investment proportion strategy is presented in more detail in Figure 2.13. The top panels show the value function and optimal investment proportion strategy for an absolute bound of $\tilde{\varepsilon}(t, x) = 0.03$ and a relative bound of $\tilde{\varepsilon}(t, x) = 0.03x$, whereas in the bottom panels the effect of an absolute bound of $\tilde{\varepsilon}(t, x) = 0.01$ and of a relative bound of $\tilde{\varepsilon}(t, x) = 0.01x$ are shown. For a small initial capital, i.e., $x < 1$, the investor who is constrained to a relative bound puts a larger proportion into both stocks compared to the investor who is constrained to an absolute bound. Thus, for $x < 1$ the associated value function of the investor who is constrained to a relative bound is above the value function of the investor who is constrained to an absolute bound. The opposite holds true for $x > 1$, whereas for $x = 1$ the value function and the investment proportion strategy are the same for the investor who is constrained to a relative bound and the investor who is constrained to an absolute bound.

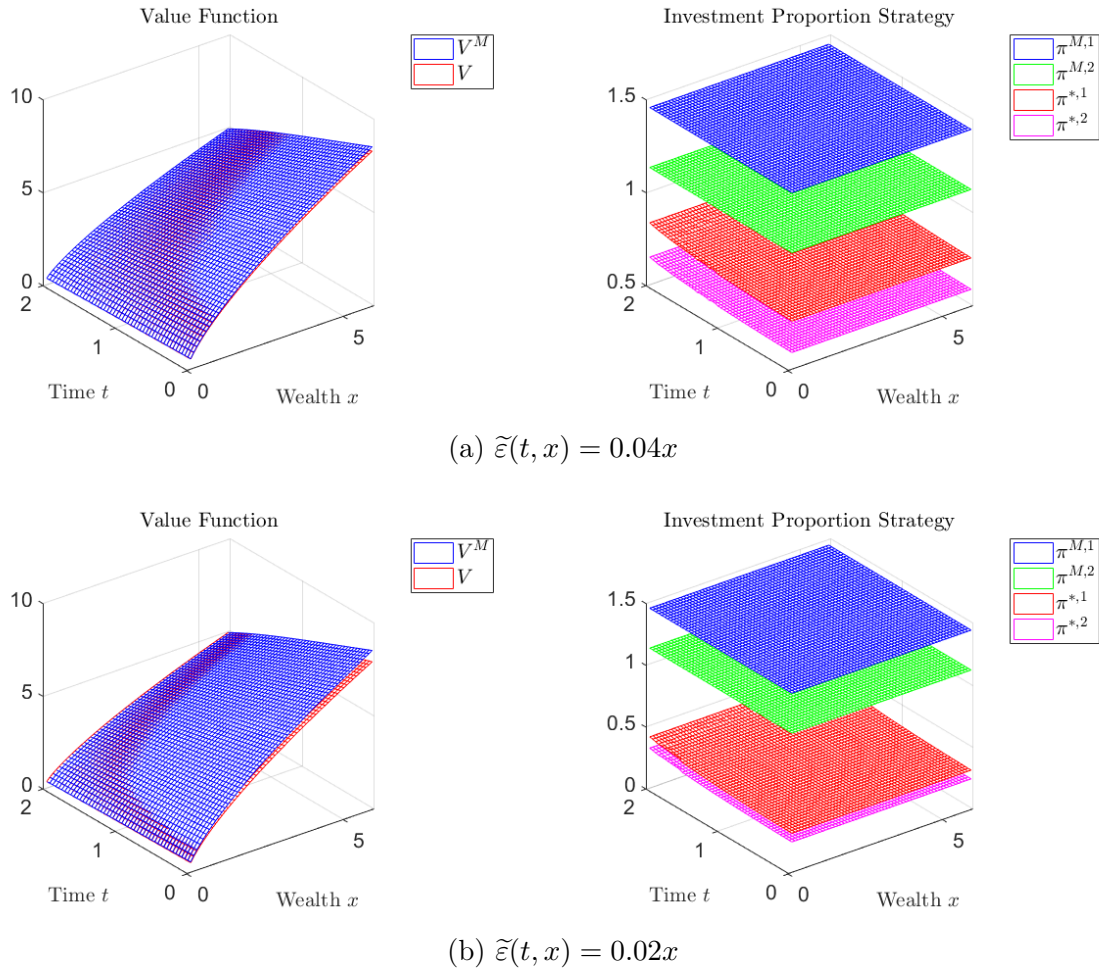
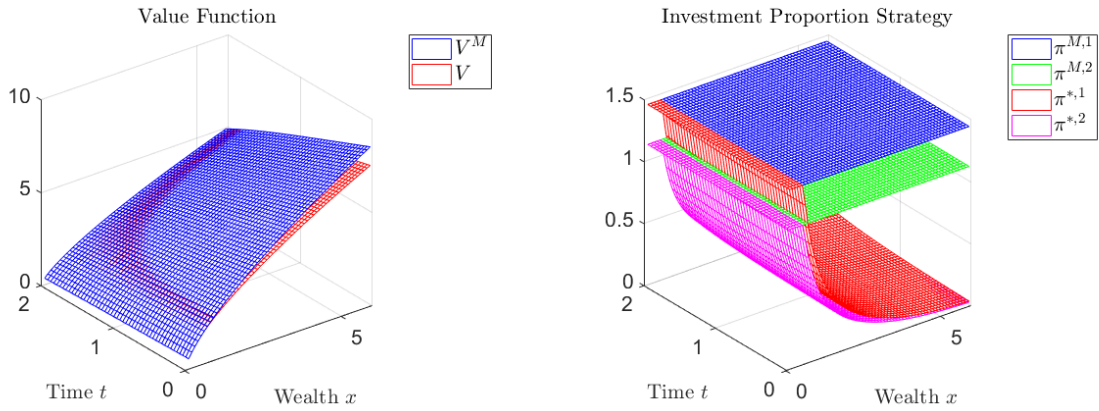
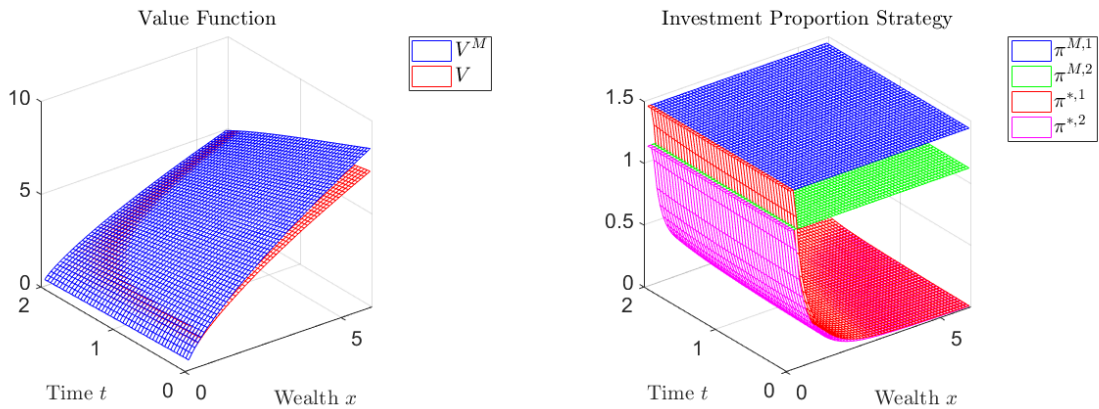


Figure 2.11: Effect of the EL constraint on the value function and optimal investment proportion strategy for different relative bounds $\tilde{\varepsilon}(t, x)$. The benchmark is $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$.



(a) $\tilde{\varepsilon}(t, x) = 0.04$



(b) $\tilde{\varepsilon}(t, x) = 0.02$

Figure 2.12: Effect of the EL constraint on the value function and optimal investment proportion strategy for different absolute bounds $\tilde{\varepsilon}(t, x)$. The benchmark is $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$.

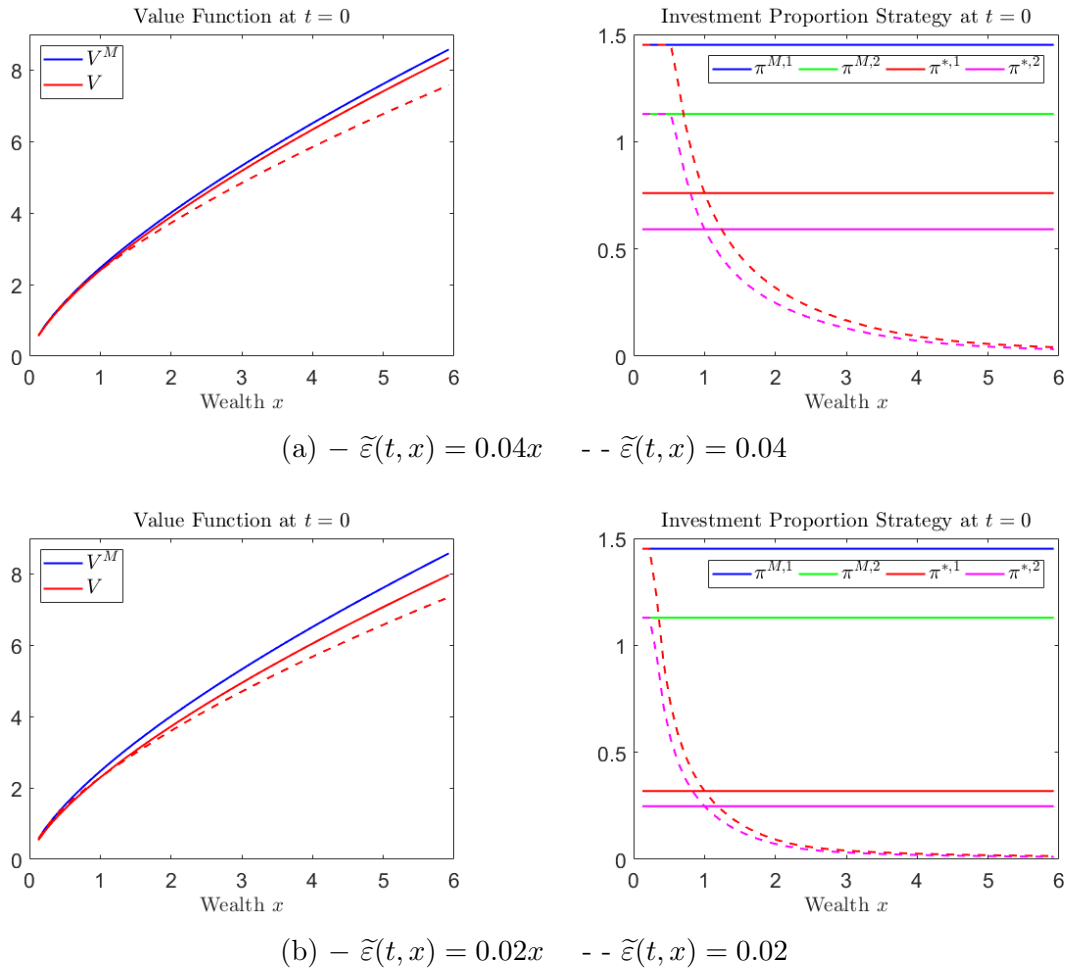


Figure 2.13: Effect of the EL constraint on the value function at $t = 0$ and optimal investment proportion strategy at $t = 0$ for different bounds $\tilde{\varepsilon}(t, x)$. The benchmark is $Y_t^M = \mathbb{E}[X_{t+\Delta}^M | X_t^{(\pi, c)}]$.

2.4 Proofs

2.4.1 Proof of Lemma 2.1.3

Under the assumption that the portfolio-proportion strategy $(\boldsymbol{\pi}, c)$ is kept constant and equal to $(\bar{\boldsymbol{\pi}}, \bar{c})$ between time t and $t + \Delta$ the wealth at time $t + \Delta$ is given by

$$X_{t+\Delta} = \exp \left\{ \ln(X_t) + \left(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} - \frac{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2}{2} \right) \Delta + \bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}(\mathbf{W}_{t+\Delta} - \mathbf{W}_t) \right\}.$$

From the equation above we obtain that $X_{t+\Delta}$ is – conditionally on \mathcal{F}_t – distributed as e^Z , where Z is normally distributed with mean m and variance s^2 . Here,

$$m := \ln(X_t) + \left(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} - \frac{1}{2}\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2 \right) \Delta \quad \text{and} \quad s^2 := \|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2 \Delta.$$

Recall, the Expected Loss at time t is defined by

$$\text{EL}_t(L_t) := \mathbb{E} \left[(Y_t - X_{t+\Delta})^+ \mid \mathcal{F}_t \right] = \mathbb{E} \left[(Y_t - e^Z)^+ \mid \mathcal{F}_t \right].$$

This expectation can be calculated as follows. Let

$$f_Z(z) = \frac{1}{\sqrt{2\pi}s} \exp \left(-\frac{(z - m)^2}{2s^2} \right)$$

denote the probability density function of Z . Then

$$\text{EL}_t(L_t) = \int_{-\infty}^{\infty} (Y_t - e^z)^+ f_Z(z) dz = \int_{-\infty}^{\ln(Y_t)} (Y_t - e^z) f_Z(z) dz = Y_t I_1 + I_2,$$

where $I_1 := \int_{-\infty}^{\ln(Y_t)} f_Z(z) dz$ and $I_2 := - \int_{-\infty}^{\ln(Y_t)} e^z f_Z(z) dz$.

For the integral I_1 an appropriate change of variables yields

$$I_1 = \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \Phi(d_1),$$

where

$$d_1 := \frac{\ln(Y_t) - m}{s} = \frac{1}{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|\sqrt{\Delta}} \left[\ln \left(\frac{\tilde{f}(t, X_t)}{X_t} \right) - \left(\bar{\boldsymbol{\pi}}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} - \frac{\|\bar{\boldsymbol{\pi}}'\boldsymbol{\sigma}\|^2}{2} \right) \Delta \right]$$

and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The integral I_2 can be written as

$$\begin{aligned} I_2 &= - \int_{-\infty}^{\ln(Y_t)} \frac{1}{\sqrt{2\pi}s} \exp \left(z - \frac{z^2 - 2zm + m^2}{2s^2} \right) dz \\ &= -e^{\frac{s^2}{2} + m} \int_{-\infty}^{\ln(Y_t)} \frac{1}{\sqrt{2\pi}s} \exp \left(-\frac{(z - (m + s^2))^2}{2s^2} \right) dz. \end{aligned}$$

Using the change of variables technique yields

$$I_2 = -e^{\frac{s^2}{2}+m} \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = -e^{\frac{s^2}{2}+m} \Phi(d_2),$$

where

$$\begin{aligned} d_2 &:= \frac{\ln(Y_t) - (m + s^2)}{s} \\ &= \frac{1}{\|\bar{\pi}'\sigma\|\sqrt{\Delta}} \left[\ln\left(\frac{\tilde{f}(t, X_t)}{X_t}\right) - \left(\bar{\pi}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c} + \frac{1}{2}\|\bar{\pi}'\sigma\|^2\right)\Delta \right]. \end{aligned}$$

Note, $\exp\{s^2/2 + m\} = X_t \exp((\bar{\pi}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c})\Delta)$. Finally we obtain

$$\text{EL}_t(L_t) = Y_t I_1 + I_2 = \tilde{f}(t, X_t) \Phi(d_1) - X_t \exp((\bar{\pi}'(\boldsymbol{\mu} - \mathbf{1}r) + r - \bar{c})\Delta) \Phi(d_2).$$

2.4.2 Proof of Lemma 2.2.1

Given an investment-consumption strategy (φ, η) the wealth process X evolves as follows

$$X_{t_{n+1}} = e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) + \varphi_{t_n} \cdot \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma(W_{t_{n+1}} - W_{t_n})\right\}.$$

Note that given a probability level $\alpha \in (0, 1)$ the Value at Risk at time t_n is defined by $\text{VaR}_{t_n}^\alpha(L_{t_n}) := \inf\{l \in \mathbb{R} \mid \mathbb{P}(L_{t_n} \geq l | \mathcal{G}_{t_n}) \leq \alpha\}$. We have

$$\begin{aligned} L_{t_n} = Y_{t_n} - X_{t_{n+1}} &= \tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) \\ &\quad - \varphi_{t_n} \cdot \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma(W_{t_{n+1}} - W_{t_n})\right\} \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}(L_{t_n} \geq l | \mathcal{G}_{t_n}) &= \mathbb{P}\left(\exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma(W_{t_{n+1}} - W_{t_n})\right\}\right. \\ &\quad \left. \leq \frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - l}{\varphi_{t_n}} \middle| \mathcal{G}_{t_n}\right) \\ &= \mathbb{P}\left(\Delta^{-\frac{1}{2}}(W_{t_{n+1}} - W_{t_n}) \leq z \middle| \mathcal{G}_{t_n}\right) = \Phi(z), \end{aligned}$$

where

$$z = \frac{1}{\sigma\sqrt{\Delta}} \left(\ln\left\{\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - l}{\varphi_{t_n}}\right\} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta \right).$$

We have used that the random variable $\Delta^{-\frac{1}{2}}(W_{t_{n+1}} - W_{t_n})$ is standard normally distributed and independent of \mathcal{G}_{t_n} . Thus, $\mathbb{P}(L_{t_n} \geq l | \mathcal{G}_{t_n}) = \Phi(z) \leq \alpha$ is satisfied for $z \leq \Phi^{-1}(\alpha)$ yielding

$$l \geq \tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - \exp\left\{\Phi^{-1}(\alpha)\sigma\sqrt{\Delta} + \left(\mu - \frac{\sigma^2}{2}\right)\Delta\right\}\varphi_{t_n}.$$

Since the Dynamic Value at Risk is the smallest l satisfying the above inequality we obtain $\text{VaR}_{t_n}^\alpha(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}, \varphi_{t_n}, \eta_{t_n})$, where

$$\tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) = \tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) - \exp\left\{\Phi^{-1}(\alpha)\sigma\sqrt{\Delta} + \left(\mu - \frac{\sigma^2}{2}\right)\Delta\right\}\bar{\varphi}.$$

The Tail Conditional Expectation at time t_n is defined by

$$\begin{aligned} \text{TCE}_{t_n}^\alpha(L_{t_n}) &= \mathbb{E}_{t_n} [L_{t_n} | L_{t_n} \geq \text{VaR}_{t_n}^\alpha(L_{t_n})] \\ &= \frac{\mathbb{E} [L_{t_n} \mathbf{I}(L_{t_n} \geq \text{VaR}_{t_n}^\alpha(L_{t_n})) | \mathcal{G}_{t_n}]}{\mathbb{P}(L_{t_n} \geq \text{VaR}_{t_n}^\alpha(L_{t_n}) | \mathcal{G}_{t_n})} \\ &= \frac{1}{\alpha} \mathbb{E} [L_{t_n} \mathbf{I}(L_{t_n} \geq \text{VaR}_{t_n}^\alpha(L_{t_n})) | \mathcal{G}_{t_n}], \end{aligned}$$

where $\mathbf{I}(A)$ denotes the indicator function of the set A . Using

$$\begin{aligned} L_{t_n} = Y_{t_n} - X_{t_{n+1}} &= \tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) \\ &\quad - \varphi_{t_n} \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma(W_{t_{n+1}} - W_{t_n})\right\} \end{aligned}$$

and

$$\begin{aligned} \text{VaR}_{t_n}^\alpha(L_{t_n}) &= \tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) \\ &\quad - \varphi_{t_n} \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta + \Phi^{-1}(\alpha)\sigma\sqrt{\Delta}\right\} \end{aligned}$$

yields that the above inequality $\text{VaR}_{t_n}^\alpha(L_{t_n}) \leq L_{t_n}$ is equivalent to $\Delta^{-\frac{1}{2}}(W_{t_{n+1}} - W_{t_n}) \leq \Phi^{-1}(\alpha)$. Thus,

$$\begin{aligned} &\mathbb{E} [L_{t_n} \mathbf{I}(L_{t_n} \geq \text{VaR}_{t_n}^\alpha(L_{t_n})) | \mathcal{G}_{t_n}] \\ &= \mathbb{E} \left[L_{t_n} \mathbf{I}\left((W_{t_{n+1}} - W_{t_n})\Delta^{-\frac{1}{2}} \leq \Phi^{-1}(\alpha)\right) \middle| \mathcal{G}_{t_n} \right] \\ &= (Y_{t_n} - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})) \mathbb{E} \left[\mathbf{I}\left((W_{t_{n+1}} - W_{t_n})\Delta^{-\frac{1}{2}} \leq \Phi^{-1}(\alpha)\right) \middle| \mathcal{G}_{t_n} \right] \\ &\quad - \varphi_{t_n} \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta\right\} \mathbb{E} \left[e^{\sigma(W_{t_{n+1}} - W_{t_n})} \right. \\ &\quad \left. \cdot \mathbf{I}\left((W_{t_{n+1}} - W_{t_n})\Delta^{-\frac{1}{2}} \leq \Phi^{-1}(\alpha)\right) \middle| \mathcal{G}_{t_n} \right] \\ &= \left(\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) \right) \alpha - \varphi_{t_n} \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\Delta\right\} \\ &\quad \cdot \mathbb{E} \left[e^{\sigma(W_{t_{n+1}} - W_{t_n})} \mathbf{I}\left((W_{t_{n+1}} - W_{t_n})\Delta^{-\frac{1}{2}} \leq \Phi^{-1}(\alpha)\right) \middle| \mathcal{G}_{t_n} \right]. \end{aligned}$$

We obtain

$$\begin{aligned} & \mathbb{E} \left[e^{\sigma(W_{t_{n+1}} - W_{t_n})} \mathbf{I} \left((W_{t_{n+1}} - W_{t_n}) \Delta^{-\frac{1}{2}} \leq \Phi^{-1}(\alpha) \right) \middle| \mathcal{G}_{t_n} \right] \\ &= \int_{-\infty}^{\Phi^{-1}(\alpha)} e^{\sigma\sqrt{\Delta}z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz, \end{aligned}$$

where we have used that the random variable $(W_{t_{n+1}} - W_{t_n}) \Delta^{-\frac{1}{2}}$ is standard normally distributed and independent of \mathcal{G}_{t_n} . We calculate the above integral by making an appropriate change of variables

$$\begin{aligned} \int_{-\infty}^{\Phi^{-1}(\alpha)} e^{\sigma\sqrt{\Delta}z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz &= e^{\frac{\sigma^2\Delta}{2}} \int_{-\infty}^{\Phi^{-1}(\alpha) - \sigma\sqrt{\Delta}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \\ &= e^{\frac{\sigma^2\Delta}{2}} \Phi \left(\Phi^{-1}(\alpha) - \sigma\sqrt{\Delta} \right). \end{aligned}$$

Finally, the Dynamic Tail Conditional Expectation can be written as

$\text{TCE}_{t_n}^\alpha(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}, \varphi_{t_n}, \eta_{t_n})$, where

$$\tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) = \tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) - \frac{1}{\alpha} e^{\mu\Delta} \Phi \left(\Phi^{-1}(\alpha) - \sigma\sqrt{\Delta} \right) \bar{\varphi}.$$

In order to prove the statement for the Dynamic Expected Loss we use that the wealth at time $X_{t_{n+1}}$ is distributed as $e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) + \varphi_{t_n} \cdot e^Z$, where Z is normally distributed with mean m and variance s^2 , where $m := (\mu - \sigma^2/2)\Delta$ and $s^2 := \sigma^2\Delta$. From the definition of the Dynamic Expected Loss we find

$$\begin{aligned} \text{EL}_{t_n}(L_{t_n}) &= \mathbb{E} \left[\left(Y_{t_n} - X_{t_{n+1}} \right)^+ \middle| \mathcal{G}_{t_n} \right] \\ &= \mathbb{E} \left[\left(\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - \varphi_{t_n} \cdot e^Z \right)^+ \middle| \mathcal{G}_{t_n} \right]. \end{aligned}$$

This expectation can be calculated as follows. Let

$$f_Z(z) = \frac{1}{\sqrt{2\pi}s} \exp \left(-\frac{(z - m)^2}{2s^2} \right)$$

denote the probability density function of Z , then

$$\begin{aligned} \text{EL}_{t_n}(L_{t_n}) &= \int_{-\infty}^{\infty} \left(\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - \varphi_{t_n} \cdot e^z \right)^+ f_Z(z) dz \\ &= \int_{-\infty}^{\tilde{d}} \left(\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) - \varphi_{t_n} \cdot e^z \right) f_Z(z) dz \\ &= \left(\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n}) \right) I_1 - \varphi_{t_n} I_2, \end{aligned}$$

where $I_1 = \int_{-\infty}^{\tilde{d}} f_Z(z) dz$, $I_2 = \int_{-\infty}^{\tilde{d}} e^z f_Z(z) dz$ and

$$\tilde{d} = \ln \left(\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})}{\varphi_{t_n}} \right).$$

The integral I_1 can be calculated by making an appropriate change of variables and we obtain

$$I_1 = \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \Phi(d_1),$$

where

$$\begin{aligned} d_1 &:= \frac{1}{s} \left(\ln \left(\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})}{\varphi_{t_n}} \right) - m \right) \\ &= \frac{1}{\sigma\sqrt{\Delta}} \left[\ln \left(\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})}{\varphi_{t_n}} \right) - \left(\mu - \frac{\sigma^2}{2} \right) \Delta \right]. \end{aligned}$$

The integral I_2 can be written as

$$\begin{aligned} I_2 &= \int_{-\infty}^{\tilde{d}} \frac{1}{\sqrt{2\pi}s} \exp \left(z - \frac{z^2 - 2zm + m^2}{2s^2} \right) dz \\ &= e^{\frac{s^2}{2} + m} \int_{-\infty}^{\tilde{d}} \frac{1}{\sqrt{2\pi}s} \exp \left(- \frac{(z - (m + s^2))^2}{2s^2} \right) dz. \end{aligned}$$

Using the change of variables method yields

$$I_2 = e^{\frac{s^2}{2} + m} \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = e^{\frac{s^2}{2} + m} \Phi(d_2),$$

where

$$\begin{aligned} d_2 &= \frac{1}{s} \left[\ln \left(\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})}{\varphi_{t_n}} \right) - (m + s^2) \right] \\ &= \frac{1}{\sigma\sqrt{\Delta}} \left[\ln \left(\frac{\tilde{f}(t_n, X_{t_n}) - e^{r\Delta}(X_{t_n} - \eta_{t_n} - \varphi_{t_n})}{\varphi_{t_n}} \right) - \left(\mu + \frac{\sigma^2}{2} \right) \Delta \right]. \end{aligned}$$

Note, $e^{s^2/2+m} = e^{\mu\Delta}$, thus we obtain $\text{EL}_{t_n}(L_{t_n}) = \tilde{\psi}(t_n, X_{t_n}, \varphi_{t_n}, \eta_{t_n})$, where

$$\tilde{\psi}(t, x, \bar{\varphi}, \bar{\eta}) = \left(\tilde{f}(t, x) - e^{r\Delta}(x - \bar{\eta} - \bar{\varphi}) \right) \Phi(d_1) - e^{\mu\Delta} \Phi(d_2) \bar{\varphi}.$$

2.5 Results from dynamic programming

In Subsection 2.5.1 and Subsection 2.5.2, we briefly introduce the dynamic programming approach for solving stochastic control problems in continuous time and discrete time, respectively. Possible references for a more comprehensive introduction to optimal stochastic control are DAVIS [25], FLEMING & RISHEL [32], FLEMING & SONER [33], KORN & KORN [51], KRYLOV [52], ØKSENDAL & SULEM [65], PHAM [67], TOUZI [79], or YONG & ZHOU [83] for continuous-time problems and BÄUERLE & RIEDER [7], BELLMAN [11], [12], [13], BOUDAREL, DELMAS & GUICHET [20], HINDERER [42], HOWARD [43], or JACOBS [44] for discrete-time problems.

2.5.1 Continuous time

We begin this subsection with formulating the stochastic control problem in continuous time by using the framework of controlled diffusion processes. Then we introduce the dynamic programming principle initiated by Bellman in the 1950s to solve the control problem. The basic idea of this approach is to embed the control problem into a family of control problems by varying the initial state values of the controlled diffusion process. It yields a nonlinear partial differential equation (PDE) of second order which is called the Hamilton-Jacobi-Bellman (HJB) equation. Supposing the existence of a regular solution to the HJB equation, its optimality is validated by the verification theorem.

We fix a finite horizon $0 < T < \infty$ and consider a control model on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ satisfying the usual conditions. The state of the control model is governed by the following stochastic differential equation (SDE) valued in \mathbb{R}^d :

$$\begin{aligned} d\mathbf{X}_s &= a(\mathbf{X}_s, \mathbf{u}_s)ds + b(\mathbf{X}_s, \mathbf{u}_s)d\mathbf{W}_s, & s \in [t, T] \\ \mathbf{X}_t &= \mathbf{x}, \end{aligned} \tag{2.21}$$

where \mathbf{W} is a m -dimensional Brownian motion and $\mathbf{u} = (\mathbf{u}_t)_{t \in [0, T]}$ is a progressively measurable (with respect to \mathbb{F}) control process, valued in $\mathcal{K} \subset \mathbb{R}^n$. The functions $a: \mathbb{R}^d \times \mathcal{K} \rightarrow \mathbb{R}^d$ and $b: \mathbb{R}^d \times \mathcal{K} \rightarrow \mathbb{R}^{d \times m}$ are measurable and satisfy a uniform Lipschitz condition, i.e., $\exists C \geq 0, \forall x, y \in \mathbb{R}^d, \forall k \in \mathcal{K}$ such that

$$\|a(x, k) - a(y, k)\| + \|b(x, k) - b(y, k)\| \leq C\|x - y\|.$$

Definition 2.5.1. The set of control processes \mathbf{u} satisfying

$$\mathbb{E} \left[\int_0^T \|a(0, \mathbf{u}_s)\|^2 + \|b(0, \mathbf{u}_s)\|^2 ds \right] < \infty$$

is denoted by \mathcal{A} .

Remark 2.5.2. For all controls $\mathbf{u} \in \mathcal{A}$ and for any initial condition $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d$, there exists a unique strong solution to the SDE (2.21) satisfying

$$\mathbb{E}_{t,\mathbf{x}} \left[\sup_{s \in [t, T]} \|\mathbf{X}_s\|^2 \right] < \infty \quad \text{and} \quad \lim_{h \downarrow 0} \mathbb{E}_{t,\mathbf{x}} \left[\sup_{s \in [t, t+h]} \|\mathbf{X}_s - \mathbf{x}\|^2 \right] = 0, \quad (2.22)$$

where the term $\mathbb{E}_{t,\mathbf{x}}[\cdot]$ denotes the conditional expectation given the information known up to time t and $\mathbf{X}_t = \mathbf{x}$. A proof can be found in KRYLOV [52, Chapter 2].

Definition 2.5.3. Let $f: [0, T] \times \mathbb{R}^d \times \mathcal{K} \rightarrow \mathbb{R}$ a measurable function. The set of control processes $\mathbf{u} \in \mathcal{A}$ satisfying

$$\mathbb{E}_{t,\mathbf{x}} \left[\int_t^T |f(s, \mathbf{X}_s, \mathbf{u}_s)| ds \right] < \infty$$

is denoted by $\mathcal{A}(t, \mathbf{x})$.

Given a control process $\mathbf{u} = (\mathbf{u}_t)_{t \in [0, T]} \in \mathcal{A}(t, \mathbf{x})$ we define the reward function $\mathcal{J}(\cdot, \cdot, \mathbf{u}): [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d$ by

$$\mathcal{J}(t, \mathbf{x}, \mathbf{u}) = \mathbb{E}_{t,\mathbf{x}} \left[\int_t^T f(s, \mathbf{X}_s, \mathbf{u}_s) ds + g(\mathbf{X}_T) \right], \quad (2.23)$$

where $g: \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function which is either lower-bounded or satisfies a quadratic growth condition for some constant C , i.e., $|g(\mathbf{x})| \leq C(1 + \|\mathbf{x}\|^2)$ for all $\mathbf{x} \in \mathbb{R}^d$. The objective is to maximize the reward function over all control processes $\mathbf{u} \in \mathcal{A}(t, \mathbf{x})$. We introduce the associated value function $V(\cdot, \cdot): [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d$ by

$$V(t, \mathbf{x}) = \sup_{\mathbf{u} \in \mathcal{A}(t, \mathbf{x})} \mathcal{J}(t, \mathbf{x}, \mathbf{u}).$$

A control process \mathbf{u}^* with $V(t, \mathbf{x}) = \mathcal{J}(t, \mathbf{x}, \mathbf{u}^*)$ is called optimal. To obtain an optimal control on the whole time interval $[t, T]$ the basic idea is to split the optimization problem into two parts. First search for an optimal control from time $\theta \in [t, T]$ given the state value \mathbf{X}_θ , i.e., compute $V(\theta, \mathbf{X}_\theta)$. Second, maximize the quantity

$$\mathbb{E}_{t,\mathbf{x}} \left[\int_t^\theta f(s, \mathbf{X}_s, \mathbf{u}_s) ds + V(\theta, \mathbf{X}_\theta) \right]$$

over controls on $[t, \theta]$. This principle is called the dynamic programming principle (DPP) and is formulated as follows:

Theorem 2.5.4 (Dynamic Programming Principle). *Let $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d$ and $\mathcal{T}_{t,T}$ the set of stopping times valued in $[t, T]$. Then it holds*

$$\begin{aligned} V(t, \mathbf{x}) &= \sup_{\mathbf{u} \in \mathcal{A}(t, \mathbf{x})} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}_{t, \mathbf{x}} \left[\int_t^\theta f(s, \mathbf{X}_s, \mathbf{u}_s) ds + V(\theta, \mathbf{X}_\theta) \right] \\ &= \sup_{\mathbf{u} \in \mathcal{A}(t, \mathbf{x})} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}_{t, \mathbf{x}} \left[\int_t^\theta f(s, \mathbf{X}_s, \mathbf{u}_s) ds + V(\theta, \mathbf{X}_\theta) \right]. \end{aligned}$$

For a proof see PHAM [67, Chapter 3, Theorem 3.3.1].

A stronger version of the DPP is written as

$$V(t, \mathbf{x}) = \sup_{\mathbf{u} \in \mathcal{A}(t, \mathbf{x})} \mathbb{E}_{t, \mathbf{x}} \left[\int_t^\theta f(s, \mathbf{X}_s, \mathbf{u}_s) ds + V(\theta, \mathbf{X}_\theta) \right], \quad (2.24)$$

for any stopping time θ valued in $[t, T]$. From the DPP we formally derive the dynamic programming equation, which is also called the Hamilton-Jacobi-Bellman (HJB) equation, by sending the stopping time θ in (2.24) to t . Hence, the HJB equation is the infinitesimal version of the DPP describing the local behavior of the value function.

Let $\theta = t + h$ and $\bar{\mathbf{u}}$ a constant control, for some arbitrary $\bar{\mathbf{u}}$ in \mathcal{K} . From the DPP (2.24) we obtain

$$V(t, \mathbf{x}) \geq \mathbb{E}_{t, \mathbf{x}} \left[\int_t^{t+h} f(s, \mathbf{X}_s, \bar{\mathbf{u}}) ds + V(t+h, \mathbf{X}_{t+h}) \right]. \quad (2.25)$$

Under the assumption that V is smooth enough, applying Itô's formula on $V(t+h, \mathbf{X}_{t+h})$ and substituting into (2.25) yields the inequality

$$0 \geq \mathbb{E}_{t, \mathbf{x}} \left[\int_t^{t+h} f(s, \mathbf{X}_s, \bar{\mathbf{u}}) ds + \int_t^{t+h} \left(\frac{\partial V}{\partial t} + \mathcal{H}^{\bar{\mathbf{u}}} V \right) (s, \mathbf{X}_s) ds \right],$$

where $\mathcal{H}^{\bar{\mathbf{u}}}$ is the operator associated to the diffusion process (2.21) for the constant control $\bar{\mathbf{u}}$ and given by

$$\mathcal{H}^{\bar{\mathbf{u}}} V = a(\mathbf{x}, \bar{\mathbf{u}}) D_{\mathbf{x}} V + \frac{1}{2} \text{tr}(b(\mathbf{x}, \bar{\mathbf{u}}) b'(\mathbf{x}, \bar{\mathbf{u}}) D_{\mathbf{x}}^2 V).$$

By dividing by h and sending h to 0, we then obtain from the mean-value theorem

$$0 \geq f(t, \mathbf{x}, \bar{\mathbf{u}}) + \frac{\partial V}{\partial t}(t, \mathbf{x}) + \mathcal{H}^{\bar{\mathbf{u}}} V(t, \mathbf{x}).$$

Since $\bar{\mathbf{u}} \in \mathcal{K}$ was chosen arbitrarily it follows

$$0 \geq \frac{\partial V}{\partial t}(t, \mathbf{x}) + \sup_{\bar{\mathbf{u}} \in \mathcal{K}} [f(t, \mathbf{x}, \bar{\mathbf{u}}) + \mathcal{H}^{\bar{\mathbf{u}}} V(t, \mathbf{x})]. \quad (2.26)$$

On the other hand, assume an optimal control \mathbf{u}^* exists and \mathbf{X}^* is the associated solution to (2.21). Then, the DPP (2.24) yields

$$V(t, \mathbf{x}) = \mathbb{E}_{t, \mathbf{x}} \left[\int_t^{t+h} f(s, \mathbf{X}_s^*, \mathbf{u}_s^*) ds + V(t+h, \mathbf{X}_{t+h}^*) \right].$$

Using similar arguments as above, we obtain the equality

$$0 = \frac{\partial V}{\partial t}(t, \mathbf{x}) + f(t, \mathbf{x}, \mathbf{u}^*) + \mathcal{H}^{\mathbf{u}^*} V(t, \mathbf{x}). \quad (2.27)$$

By combining (2.26) and (2.27) it follows

$$0 = \frac{\partial V}{\partial t}(t, \mathbf{x}) + \sup_{\bar{\mathbf{u}} \in \mathcal{K}} [f(t, \mathbf{x}, \bar{\mathbf{u}}) + \mathcal{H}^{\bar{\mathbf{u}}} V(t, \mathbf{x})], \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d, \quad (2.28)$$

if the above supremum in $\bar{\mathbf{u}}$ is finite. The equation (2.28) is called the dynamic programming equation or Hamilton-Jacobi-Bellman (HJB) equation. We immediately obtain the terminal condition associated to this PDE from the definition (2.23) of the reward function considered at the horizon date T , i.e.,

$$V(T, \mathbf{x}) = g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

To prove that, given a smooth solution to the HJB equation, this candidate coincides with the value function we formulate the verification theorem, which is based essentially on Itô's formula.

Theorem 2.5.5 (Verification Theorem). *Let w be a function in $C^{1,2}([0, T] \times \mathbb{R}^d) \cap C^0([0, T] \times \mathbb{R}^d)$, satisfying the quadratic growth condition*

$$|w(t, \mathbf{x})| \leq C(1 + \|\mathbf{x}\|^2), \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d,$$

for some constant C . Suppose that w is a solution to the HJB equation

$$\begin{aligned} \frac{\partial w}{\partial t}(t, \mathbf{x}) + \sup_{\bar{\mathbf{u}} \in \mathcal{K}} [f(t, \mathbf{x}, \bar{\mathbf{u}}) + \mathcal{H}^{\bar{\mathbf{u}}} w(t, \mathbf{x})] &= 0, \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d, \\ w(T, \mathbf{x}) &= g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d. \end{aligned} \quad (2.29)$$

Suppose further that there exists a measurable function $\hat{\mathbf{u}}(t, \mathbf{x}): [0, T] \times \mathbb{R}^d \rightarrow \mathcal{K}$ such that

- the SDE

$$\begin{aligned} d\mathbf{X}_s &= a(\mathbf{X}_s, \hat{\mathbf{u}}(s, \mathbf{X}_s)) ds + b(\mathbf{X}_s, \hat{\mathbf{u}}(s, \mathbf{X}_s)) d\mathbf{W}_s, \quad s \in (t, T] \\ \mathbf{X}_t &= \mathbf{x}, \end{aligned}$$

admits a unique solution $\hat{\mathbf{X}}$,

- the process $\{\hat{\mathbf{u}}(s, \hat{\mathbf{X}}_s): t \leq s \leq T\}$ lies in $\mathcal{A}(t, \mathbf{x})$ and
- the following holds

$$\hat{\mathbf{u}}(t, \mathbf{x}) \in \arg \max_{\bar{\mathbf{u}} \in \mathcal{K}} [f(t, \mathbf{x}, \bar{\mathbf{u}}) + \mathcal{H}^{\bar{\mathbf{u}}} w(t, \mathbf{x})].$$

Then w coincides with the value function, i.e., $w(t, \mathbf{x}) = V(t, \mathbf{x})$ on $[0, T] \times \mathbb{R}^d$, and $\hat{\mathbf{u}}$ is an optimal control.

Proof. Let $\mathbf{u} \in \mathcal{A}(t, \mathbf{x})$ an arbitrary control and τ a stopping time valued in $[t, \infty)$. Since $w \in C^{1,2}([0, T] \times \mathbb{R}^d)$ we can apply Itô's formula yielding

$$\begin{aligned} w(s \wedge \tau, \mathbf{X}_{s \wedge \tau}) = & w(t, \mathbf{x}) + \int_t^{s \wedge \tau} \frac{\partial w}{\partial t}(z, \mathbf{X}_z) + \mathcal{H}^{\mathbf{u}_z}(z, \mathbf{X}_z) dz \\ & + \int_t^{s \wedge \tau} D_x w(z, \mathbf{X}_z)' b(\mathbf{X}_z, \mathbf{u}_z) d\mathbf{W}_z, \end{aligned} \quad (2.30)$$

for all $s \in [t, T)$. We choose the stopping time

$$\tau = \tau_n = \inf \left\{ s \geq t: \int_t^s |D_x w(z, \mathbf{X}_z)' b(\mathbf{X}_z, \mathbf{u}_z) d\mathbf{W}_z|^2 \geq n \right\},$$

and obtain $\tau_n \nearrow \infty$ when $n \rightarrow \infty$. Hence, the stopped process

$$\int_t^{s \wedge \tau_n} D_x w(z, \mathbf{X}_z)' b(\mathbf{X}_z, \mathbf{u}_z) d\mathbf{W}_z$$

is a martingale and it holds

$$\mathbb{E}_{t, \mathbf{x}} \left[\int_t^{s \wedge \tau_n} D_x w(z, \mathbf{X}_z)' b(\mathbf{X}_z, \mathbf{u}_z) d\mathbf{W}_z \right] = 0.$$

By taking the expectation in (2.30) it follows

$$\mathbb{E}_{t, \mathbf{x}} [w(s \wedge \tau_n, \mathbf{X}_{s \wedge \tau_n})] = w(t, \mathbf{x}) + \mathbb{E}_{t, \mathbf{x}} \left[\int_t^{s \wedge \tau_n} \frac{\partial w}{\partial t}(z, \mathbf{X}_z) + \mathcal{H}^{\mathbf{u}_z}(z, \mathbf{X}_z) dz \right].$$

Since w is a solution to the HJB equation (2.29) we obtain

$$\frac{\partial w}{\partial t}(z, \mathbf{X}_z) + \mathcal{H}^{\mathbf{u}_z}(z, \mathbf{X}_z) + f(z, \mathbf{X}_z, \mathbf{u}_z) \leq 0.$$

Thus,

$$\mathbb{E}_{t,\mathbf{x}} [w(s \wedge \tau_n, \mathbf{X}_{s \wedge \tau_n})] \leq w(t, \mathbf{x}) - \mathbb{E}_{t,\mathbf{x}} \left[\int_t^{s \wedge \tau_n} f(z, \mathbf{X}_z, \mathbf{u}_z) dz \right]. \quad (2.31)$$

Furthermore, it holds

$$\left| \int_t^{s \wedge \tau_n} f(z, \mathbf{X}_z, \mathbf{u}_z) dz \right| \leq \int_t^{s \wedge \tau_n} |f(z, \mathbf{X}_z, \mathbf{u}_z)| dz \leq \int_t^T |f(z, \mathbf{X}_z, \mathbf{u}_z)| dz,$$

where the right-hand-side term is integrable by the integrability condition on $\mathcal{A}(t, \mathbf{x})$. Since w satisfies a quadratic growth condition, it follows

$$w(s \wedge \tau_n, \mathbf{X}_{s \wedge \tau_n}) \leq C (1 + \|\mathbf{X}_{s \wedge \tau_n}\|^2) \leq C \left(1 + \sup_{s \in [t, T]} \|\mathbf{X}_s\|^2 \right),$$

where the right-hand-side term is integrable from (2.22). We can then send n to infinity in (2.31) and the dominated convergence theorem yields

$$\mathbb{E}_{t,\mathbf{x}} [w(s, \mathbf{X}_s)] \leq w(t, \mathbf{x}) - \mathbb{E}_{t,\mathbf{x}} \left[\int_t^s f(z, \mathbf{X}_z, \mathbf{u}_z) dz \right]. \quad (2.32)$$

Since w is continuous and by (2.29), we have

$$\lim_{s \rightarrow T} w(s, \mathbf{X}_s) = w(T, \mathbf{X}_T) = g(\mathbf{X}_T).$$

Thus, sending s to T into (2.32) yields by the dominated convergence theorem

$$w(t, \mathbf{x}) \geq \mathbb{E}_{t,\mathbf{x}} \left[\int_t^T f(z, \mathbf{X}_z, \mathbf{u}_z) dz + g(\mathbf{X}_T) \right] = \mathcal{J}(t, \mathbf{x}, \mathbf{u}).$$

From the arbitrariness of $\mathbf{u} \in \mathcal{A}(t, \mathbf{x})$, it follows

$$w(t, \mathbf{x}) \geq \sup_{\mathbf{u} \in \mathcal{A}(t, \mathbf{x})} \mathcal{J}(t, \mathbf{x}, \mathbf{u}) = V(t, \mathbf{x}), \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d. \quad (2.33)$$

On the other hand, by definition of the control process $\{\hat{\mathbf{u}}(s, \mathbf{X}_s) : t \leq s \leq T\}$, we have

$$\frac{\partial w}{\partial t}(t, \mathbf{x}) + \mathcal{H}^{\hat{\mathbf{u}}(t, \mathbf{x})}(t, \mathbf{x}) + f(t, \mathbf{x}, \hat{\mathbf{u}}(t, \mathbf{x})) = 0, \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^d.$$

Using similar arguments as above, we obtain

$$w(t, \mathbf{x}) = \mathcal{J}(t, \mathbf{x}, \hat{\mathbf{u}}) \leq V(t, \mathbf{x}). \quad (2.34)$$

Combining (2.33) and (2.34) finally shows that $w = V$ with $\hat{\mathbf{u}}$ as an optimal control. \square

Remark 2.5.6. As a byproduct, we obtain by the verification theorem that the optimal control process \mathbf{u}^* is in the form $\mathbf{u}_s^* = u(s, \mathbf{X}_s^*)$ for some measurable function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathcal{K}$. Such controls are called Markovian controls.

Remark 2.5.7. The existence of a smooth solution to the HJB equation can be shown using the existence results provided in FLEMING & RISHEL [32, Chapter VI, Theorem 6.1] and FLEMING & SONER [33, Chapter IV, Theorem 4.2]. In the particular case where the required conditions of sufficiently smooth and bounded functions a , b , f , and g are not fulfilled the problem can be tackled by considering a sequence of control problems, where the discontinuous functions are approximated by mollified functions as proposed in FLEMING & SONER [33, Appendix C]. Furthermore, the unbounded functions are suitably truncated to bounded functions using the technique described in HARRISON & ZEEVI [40, Appendix B].

2.5.2 Discrete time

In this subsection we consider a Markov Decision Process with a finite time horizon. For the sake of a short exposition of the theory of Markov Decision Problems we restrict ourselves to the case of finite state and action spaces. We show that the associated optimization problem can be solved by a backward induction algorithm. For the theory of Markov Decision Problems with general state and action spaces we refer to BÄUERLE & RIEDER [7]. The authors impose a so-called structure assumption on the problem which is often easily verified in applications and is used to prove the validity of the backward induction algorithm and the existence of optimal strategies.

Definition 2.5.8. A Markov Decision Process with planning horizon $N \in \mathbb{N}_+$ consists of a set of data (X, A, p, r, g) with the following meaning for all $n = 0, 1, \dots, N - 1$:

- X is the finite state space.
- A is the finite action space.
- $p_n(y|x, a)$ describes the transition probability, i.e., for any $x, y \in X$, $a \in A$

$$p_n(y|x, a) = \mathbb{P}(X_{n+1} = y | X_n = x, A_n = a)$$

is the probability that the state at time $n + 1$ is y given the current state is x and action a is taken.

- $r_n(x, a, y)$ is the reward obtained when taking action a at time n and a transition from a state x at time n to a state y at time $n + 1$ is observed.
- $g_N(x)$ gives the terminal reward at time N if the state is x .

Remark 2.5.9. In what follows we will often use $r_n(x, a)$, which can be considered as the expected value of $r_n(x, a, X_{n+1})$, i.e.,

$$r_n(x, a) = \mathbb{E}[r_n(x, a, X_{n+1}) | X_n = x].$$

Definition 2.5.10. For $n = 0, 1, \dots, N-1$ a decision rule at time n is denoted by u_n and defined as a mapping from the state space to the action space, i.e., $u_n: X \rightarrow A$ and $u_n(x)$ denotes the action chosen in x at time n . A sequence of decision rules $\mathbf{u} = (u_0, u_1, \dots, u_{N-1})$ is called a strategy.

Given a strategy $\mathbf{u} = (u_0, u_1, \dots, u_{N-1})$ we define the reward functional $\mathcal{J}(n, x, \mathbf{u})$ for all $n = 0, 1, \dots, N$ and $x \in X$ by

$$\mathcal{J}(n, x, \mathbf{u}) = \mathbb{E} \left[\sum_{m=n}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) | X_n = x; \mathbf{u} \right].$$

The objective is to maximize the reward function over all strategies \mathbf{u} . We introduce the associated value function for all $n = 0, 1, \dots, N$ and $x \in X$ by

$$V(n, x) = \sup_{\mathbf{u} \in \mathcal{U}} \mathcal{J}(n, x, \mathbf{u}),$$

where \mathcal{U} is some strategy set of interest. A strategy $\mathbf{u}^* = (u_0^*, u_1^*, \dots, u_{N-1}^*) \in \mathcal{U}$ with $V(0, x) = \mathcal{J}(0, x, \mathbf{u}^*)$ is called optimal.

In the next theorem we show that for a fixed strategy $\mathbf{u} = (u_0, u_1, \dots, u_{N-1})$ we can compute the reward functional recursively by a backward induction algorithm.

Theorem 2.5.11. Let $\mathbf{u} = (u_0, u_1, \dots, u_{N-1})$ be a fixed strategy. For a time $n = 0, 1, \dots, N-1$ and a state $x \in X$ the reward functional satisfies

$$\begin{aligned} \mathcal{J}(N, x, \mathbf{u}) &= g_N(x) \\ \mathcal{J}(n, x, \mathbf{u}) &= r_n(x, u_n(x)) + \sum_{y \in X} p_n(y|x, u_n(x)) \mathcal{J}(n+1, y, \mathbf{u}). \end{aligned}$$

Proof. We have

$$\begin{aligned}
\mathcal{J}(n, x, \mathbf{u}) &= \mathbb{E} \left[\sum_{m=n}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) \mid X_n = x; \mathbf{u} \right] \\
&= \mathbb{E} \left[r_n(X_n, u_n(X_n)) \mid X_n = x; \mathbf{u} \right] \\
&\quad + \mathbb{E} \left[\sum_{m=n+1}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) \mid X_n = x; \mathbf{u} \right] \\
&= r_n(x, u_n(x)) \\
&\quad + \mathbb{E} \left[\mathbb{E} \left[\sum_{m=n+1}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) \mid X_{n+1}; X_n = x; \mathbf{u} \right] \mid X_n = x; \mathbf{u} \right] \\
&= r_n(x, u_n(x)) \\
&\quad + \sum_{y \in X} p_n(y \mid x, u_n(x)) \mathbb{E} \left[\sum_{m=n+1}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) \mid X_{n+1} = y; \mathbf{u} \right] \\
&= r_n(x, u_n(x)) + \sum_{y \in X} p_n(y \mid x, u_n(x)) \mathcal{J}(n+1, y, \mathbf{u}),
\end{aligned}$$

where we have used the tower rule in the third and fourth equation. \square

Next we show that the value function is a solution to the so-called Bellman equation.

Theorem 2.5.12 (Bellman Equation). *The value function V satisfies the Bellman equation*

$$\begin{aligned}
V(N, x) &= g_N(x), \quad x \in X \\
V(n, x) &= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y \mid x, a) V(n+1, y) \right], \quad x \in X; n = 0, 1, \dots, N-1.
\end{aligned}$$

Proof. By definition it holds $V(N, x) = g_N(x)$ for all $x \in X$. We write a strategy $\mathbf{u} = (u_n, u_{n+1}, \dots, u_{N-1})$ applied at state x from time $n = 0, 1, \dots, N-1$ on as $\mathbf{u} = (a, \bar{\mathbf{u}})$ with $a \in A$ is the action chosen in x at time n and $\bar{\mathbf{u}} = (u_{n+1}, u_{n+2}, \dots, u_{N-1})$.

It follows

$$\begin{aligned}
V(n, x) &= \sup_{\mathbf{u}} \left[\mathbb{E} \left[\sum_{m=n}^{N-1} r_m(X_m, u_m(X_m)) + g_N(X_N) \mid X_n = x; \mathbf{u} \right] \right] \\
&= \sup_{(a, \bar{\mathbf{u}})} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) \mathcal{J}(n+1, y, \bar{\mathbf{u}}) \right] \\
&= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) \sup_{\bar{\mathbf{u}}} \mathcal{J}(n+1, y, \bar{\mathbf{u}}) \right] \\
&= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) V(n+1, y) \right],
\end{aligned}$$

where we have used Theorem 2.5.11 in the second equation. \square

In the following we show that a solution of the Bellman equation coincides with the value function and that the associated sequence of maximizers is an optimal strategy.

Theorem 2.5.13 (Verification Theorem). *For a time $n = 0, 1, \dots, N-1$ and a state $x \in X$ let w be a function satisfying the Bellman equation*

$$\begin{aligned}
w(N, x) &= g_N(x) \\
w(n, x) &= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) w(n+1, y) \right]
\end{aligned}$$

and for a time $n = 0, 1, \dots, N-1$ let $\hat{u}_n: X \rightarrow A$ be a function such that

$$\hat{u}_n(x) \in \arg \max_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) w(n+1, y) \right].$$

Then it holds $w(n, x) = V(n, x)$ for all $n = 0, 1, \dots, N$ and $x \in X$ and $\hat{\mathbf{u}} = (\hat{u}_0, \hat{u}_1, \dots, \hat{u}_{N-1})$ is an optimal strategy.

Proof. We show recursively that $w(n, x) \geq V(n, x)$. For $n = N$ we have $w(N, x) = g_N(x) = V(N, x)$ by definition. Suppose that $w(n+1, x) \geq V(n+1, x)$ for all $x \in X$. Using Theorem 2.5.12 we obtain

$$\begin{aligned}
w(n, x) &= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) w(n+1, y) \right] \\
&\geq \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) V(n+1, y) \right] = V(n, x).
\end{aligned}$$

On the other hand, we can show recursively that $w(n, x) \leq \mathcal{J}(n, x, \hat{\mathbf{u}})$. For $n = N$ we have $w(N, x) = g_N(x) = V(N, x) = \mathcal{J}(N, x, \hat{\mathbf{u}})$. Suppose that $w(n + 1, x) \leq \mathcal{J}(n + 1, x, \hat{\mathbf{u}})$ then

$$\begin{aligned} w(n, x) &= \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) w(n + 1, y) \right] \\ &\leq \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) \mathcal{J}(n + 1, y, \hat{\mathbf{u}}) \right] \\ &\leq \sup_{a \in A} \left[r_n(x, a) + \sum_{y \in X} p_n(y|x, a) V(n + 1, y) \right] = V(n, x) \leq \mathcal{J}(n, x, \hat{\mathbf{u}}). \end{aligned}$$

Hence we have $V(n, x) \leq w(n, x) \leq \mathcal{J}(n, x, \hat{\mathbf{u}}) \leq V(n, x)$ yielding $w = V$ and $\hat{\mathbf{u}}$ is an optimal strategy. \square

Chapter 3

Credit risk with asymmetric information and a switching barrier

Credit risk, or default risk, is the risk that a financial loss will be incurred if a counterparty does not fulfill its contractually agreed financial obligations in a timely manner. Quantitative credit risk models for measuring, monitoring, and managing credit risk have become central in today's complex financial industry. In Section 3.1 we review the two classical types of modeling approaches and present a very recent stream of hybrid models of the traditional approaches that tries to incorporate the best features of both model types. Then we proceed by defining a new hybrid model of credit risk in Section 3.2. This model is theoretically analyzed in Section 3.3 and illustrated by numerical examples in Section 3.4. The chapter concludes, in Section 3.5, with proofs omitted from the main text.

3.1 Credit risk models

There are two classical types of modeling approaches: the structural one considered by BLACK & SCHOLES [15], MERTON [61], and BLACK & COX [14], among others, and the reduced-form one considered by JARROW & TURNBULL [46], ARTZNER & DELBAEN [3], and DUFFIE & SINGLETON [28], among others. The key distinction between these two model types can be characterized in terms of the information set available to the modeler (see JARROW & PROTTER [45]). Structural models determine the time of default by using the evolution of the value of a firm's assets. Hence, explicit assumptions about the dynamics of firm's assets and the situation that will trigger the default event are made. This implicitly assumes that the modeler has complete knowledge of a firm's assets and liabilities, i.e., the modeler is assumed to have the same information as a firm's management. Whereas structural models provide a relationship between default risk and capital structure, reduced-form models treat the dynamics of default as an exogenous event. This implies knowledge of a

less detailed information set, i.e., the modeler has incomplete knowledge of a firm's condition, akin to public (market) information. The usage of reduced-form models is supported by the consensus in the credit risk literature that a firm's asset value process is unobservable by the market (see DUAN [26], ERICSSON & RENEBY [31], [30]). However, structural models can be transformed into reduced-form models by changing the information set to a less refined one. Incomplete information models of credit risk were proposed by DUFFIE & LANDO [27], GOLDBERG & GIESECKE [37], and ÇETIN ET AL. [22], among others.

3.1.1 Basic definitions and notation

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a given complete probability space endowed with the filtration $\mathbb{A} = (\mathcal{A}_t)_{t \geq 0}$ satisfying the usual conditions, i.e., \mathcal{A}_0 contains all the \mathbb{P} -null sets of \mathcal{A} and $\mathcal{A}_t = \bigcap_{u > t} \mathcal{A}_u$ for all $0 \leq t < \infty$.

Recall, a random variable $\tau: \Omega \rightarrow [0, \infty]$ is called a \mathbb{A} -stopping time if

$$\{\omega: \tau(\omega) \leq t\} \in \mathcal{A}_t$$

holds for every $0 \leq t \leq \infty$. Important types of stopping times are predictable times and totally inaccessible times (see PROTTER [70]).

Definition 3.1.1. A stopping time τ is called predictable, if there exists an increasing sequence of stopping times $(\tau_n)_{n \geq 1}$ such that $\tau_n < \tau$ on $\{\tau > 0\}$ for all n and $\lim_{n \rightarrow \infty} \tau_n = \tau$ a.s.

Remark 3.1.2. A predictable stopping time is *announced* by an increasing sequence of stopping times and thus, known to occur just before it happens.

Definition 3.1.3. A stopping time is called accessible, if there exists a sequence of predictable stopping times $(\tau_n)_{n \geq 1}$ such that

$$\mathbb{P}(\bigcup_{n=1}^{\infty} \{\omega: \tau_n(\omega) = \tau(\omega) < \infty\}) = \mathbb{P}(\tau < \infty).$$

Definition 3.1.4. A stopping time τ is called totally inaccessible, if for every predictable stopping time S

$$\mathbb{P}(\{\omega: \tau(\omega) = S(\omega) < \infty\}) = 0.$$

Remark 3.1.5. In contrast to a predictable stopping time is a totally inaccessible stopping time never announced by an increasing sequence of stopping times and thus, a “total surprise”.

On $(\Omega, \mathcal{A}, \mathbb{P})$, we consider a generic firm whose asset value process is denoted by $X = (X_t)_{t \geq 0}$, i.e., X_t represents the firm's asset value at time t . The credit risk of the firm is evaluated from the perspective of the modeler. The σ -algebra \mathcal{A} on Ω

describes the total information of the market and the σ -algebra $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ represents the information available to the modeler.

In a classical structural model the firm's asset value is assumed to follow a geometric Brownian motion and the modeler is assumed to know at time t the firm's asset value X_t , i.e., the modeler is assumed to observe the information provided by the σ -algebra $\mathbb{G} = (\sigma(X_s : 0 \leq s \leq t))_{t \geq 0}$. The firm defaults when its asset value crosses below some threshold L representing the firm's liabilities. Then the default time, denoted by τ , is given by

$$\tau = \inf\{t > 0 : X_t \leq L\}.$$

Here, the default time is a predictable stopping time for the filtration \mathbb{G} , since the sequence $\tau_n = \inf\{t > 0 : X_t \leq L + \frac{1}{n}\}$ is an announcing sequence of τ . Assuming that the firm's only liability consists of a zero-coupon bond promising to pay the face value 1 at its maturity time $T \in (0, \infty)$, if no default occurs by T and zero otherwise. The credit yield spread $S(t, T)$ on this zero-coupon bond issued by the firm is the difference between the yield of the zero-coupon bond and that of a credit risk-free zero-coupon bond maturing at time T . The short-term credit spread at time t is the credit spread for maturity T approaching t , i.e.,

$$\lim_{T \downarrow t} S(t, T),$$

and represents the excess yield over the risk-free yield over an infinitesimal time period. An immutable consequence of the predictability of the default is that short-term credit spreads forecasts of the structural model are zero implying that over a short term, a junk bond is not riskier than a bond not exposed to default risk (see GOLDBERG & GIESECKE [37]). In reality, however, default often comes unexpectedly and short-term credit spreads are positive.

The difficulty of fitting short-term credit spreads is addressed by the reduced-form approach, where default occurs unexpectedly as a true surprise. The simplest reduced-form model assumes that the default time is a stopping time generated by an exogenously given Cox process $(N_t)_{t \geq 0}$ given by $N_t = \mathbb{1}_{\{\tau \leq t\}}$ with an intensity process $(\lambda_t)_{t \geq 0}$ depending on a vector of state variables Y_t that are in the information set of the market. The modeler is assumed to observe the filtration generated by the default time τ and the vector of state variables Y_t , i.e., $\mathcal{G}_t = \sigma(\tau, Y_s : s \leq t)$. The intensity λ_t expresses the conditional default arrival rate at time t , i.e.,

$$\lambda_t = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(\tau \leq t + h | \mathcal{G}_t) \quad \text{a.s.,} \quad t < \tau.$$

Hence, default probabilities can be easily calculated from the intensity. Moreover, short-term credit spreads are directly given by the intensity (see LANDO [54]), i.e.,

$$\lim_{T \downarrow t} S(t, T) = \lambda_t \quad \text{a.s.,} \quad t < \tau.$$

A successful transformation of a structural model into a reduced-form model to gain both the economic appeal of the first and the empirical plausibility and the tractability of the latter is based on incomplete information (see GOLDBERG & GIESECKE [37]). A variant of introducing noise to the classical structural model was first proposed by DUFFIE & LANDO [27], where it is assumed that the modeler observes the firm's asset value not continuously, and receives instead only imperfect information about the firm's value at discrete time intervals. This accounts for the fact that the modeler may not have an accurate picture of a firm's asset value. More precisely, the modeler obtains at times $T_1 < T_2 < \dots < T_J$, $J \in \mathbb{N}$, a noisy report of the firm's asset value, i.e.,

$$Z_{T_j} = X_{T_j} + Y_{T_j}, \quad j \in \{1, \dots, J\},$$

where Y_{T_j} is some independent random variable. The variance of Y_{T_j} represents a measure of the degree of noise of the firm's asset value at time T_j . Since the modeler observes Z and not X the modeler's information at time t is given by the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ defined by

$$\mathcal{G}_t = \sigma(Z_s : s \leq t, s \in \{T_1, \dots, T_J\}).$$

A measure of the firm's credit quality is the conditional survival probability

$$\mathbb{P}(\tau > u | \mathcal{G}_t), \quad u \geq t.$$

Duffie and Lando were able to show that τ is a totally inaccessible stopping time for the filtration \mathbb{G} admitting an intensity λ with

$$\mathbb{1}_{\{t \geq \tau\}} - \int_0^t \lambda_s ds$$

being a local martingale. Here, the structural model is transformed into an intensity-based or reduced-form model due to information reduction and obscuring such that between observation times of the firm's asset value, the modeler does not know how the asset value has evolved and default could occur unexpectedly prior to the next observation.

Another approach to introduce noise to the system is that proposed by GOLDBERG & GIESECKE [37], where it is assumed that the modeler continuously observes the firm's asset value process but not the default barrier. The default barrier is modeled by a random variable. This model accounts for the fact that the modeler may not have an accurate picture of a firm's liabilities. Here, the default time τ is rendered totally inaccessible since it depends on the constant default barrier L that cannot be observed by the modeler. The short-term credit spread at time t is non-zero, if the firm's asset value X_t is at its historical low and zero, if the firm's asset value X_t is above its historical low.

3.2 Model for the default event

In this section a structural model is introduced, where short-term default risk is included by making the default barrier unobservable. The usual assumption in a structural model of a constant default barrier is relaxed to allow the firm's management to adjust the barrier. In Section 3.2.1 the dynamics of the firm's assets are modeled by a geometric Brownian motion and the default barrier and its adjustment time points are specified. Section 3.2.2 deals with different information structures. After the available information of the management is described, two information structures for public bond investors are distinguished. The first type of investors has complete information about the firm's assets while the second type obtains information on the firm's assets only at discrete dates. The first information structure is a standard one in classical structural models. However, the second information structure is realistic since investors usually observe the value of the firm's assets only at the times of corporate news release and not continuously.

3.2.1 Default barrier

The default event is specified in terms of the firm's asset value process $(X_t)_{t \geq 0}$ and the default threshold L . Default occurs when the value of the firm decreases to the level of the default barrier for the first time. The filtration generated by the asset process X is denoted by $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and assumed to satisfy the usual conditions, i.e., $\mathcal{F}_t = \sigma(X_s: s \leq t) \vee \mathcal{N}$, where \mathcal{N} denotes the \mathbb{P} -null sets. We model the value X of the firm's assets by a geometric Brownian motion, i.e.,

$$dX_t = X_t(\mu dt + \sigma dB_t), \quad X_0 = x_0, \quad (3.1)$$

where $(B_t)_{t \geq 0}$ is a \mathbb{F} -Brownian motion, $\mu \in \mathbb{R}$ and $\sigma > 0$. The solution to (3.1) is known to be

$$X_t = x_0 e^{mt + \sigma B_t}, \quad (3.2)$$

where $m = \mu - \sigma^2/2$. For the sake of easier notation we assume w.l.o.g. $x_0 = 1$, i.e., we assume that the asset process starts at 1. We say that a firm defaults when it stops fulfilling a contractual commitment to meet its obligations stated in a financial contract. The firm's management decides whether and when to default. Thus, the management determines the default triggering barrier. The essential difference from a classical structural model is that the management is not constrained to decide on one fixed barrier but it can dynamically adjust the default barrier. This reflects the management's possibility to react to changes in the economic environment or the election of a new firm management. The time points at which the management adjusts the default barrier are deterministic and denoted by t_i , $i = 0, \dots, n-1$, with

$0 = t_0 < t_1 < \dots < t_{n-1} < T$, where $T \in (0, \infty)$ is a finite time horizon. Further, we denote $T =: t_n$. The default barrier L can be written as

$$L_t = \sum_{i=1}^n L^i \mathbb{1}_{[t_{i-1}, t_i)},$$

where L^i , $i = 1, \dots, n$ are \mathcal{A} -measurable random variables representing the private information of the management on the default barrier and $\mathbb{1}_A$ denotes the indicator function of the set A . Public investors do not have any knowledge on the default barrier except that they know the adjustment time points t_0, \dots, t_{n-1} and they agree on the joint probability distributions for L^1, \dots, L^i , $i = 1, \dots, n$, which are denoted by F_{L^1, \dots, L^i} . The associated probability density functions are denoted by f_{L^1, \dots, L^i} for $i = 1, \dots, n$. We make the assumption that $\mathbf{L} = (L^1, \dots, L^n)$ is independent of \mathcal{F}_T . The random default time is given by $\tau = \inf\{t > 0: X_t \leq L_t\}$. The running minimum asset process is denoted by M and given by $M_t = \inf\{X_u: 0 \leq u < t\}$. Further, the running minimum started from a certain time point s is denoted by $M_{[s,t)} = \inf\{X_u: s \leq u < t\}$. For $s = 0$ we have $M_{[0,t)} = M_t$. This model setup is illustrated by Figure 3.1, which shows two trajectories of the firm value X , $n = 3$ default barriers L^1 , L^2 , and L^3 , and the three running minimum processes $M_{[t_{i-1}, t_i)}$, for $i = 1, 2, 3$. We consider two scenarios. In the first scenario shown in the left panel there occurs no default whereas in the second scenario shown in the right panel default occurs between t_2 and $t_3 = T$.

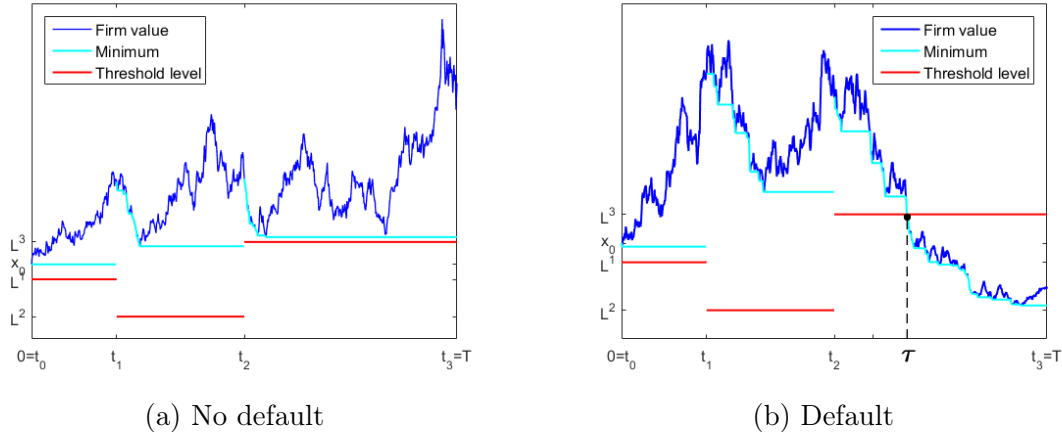


Figure 3.1: Plot of two trajectories of the firm's asset process X_t , the associated running minimum process $M_{[t_{i-1}, t)}$ and the default thresholds L^i .

3.2.2 Information structure

We distinguish the following three information structures.

Management's information

The management has complete information about the firm's asset process X and obtains information on the default threshold L^i at time t_{i-1} . Thus, the management's information structure $\mathbb{G}^M = (\mathcal{G}_t^M)_{t \in [0, T]}$ is a progressive enlargement of the filtration \mathbb{F} by the default threshold process $(L_t)_{t \geq 0}$, i.e.,

$$\mathcal{G}_t^M = \mathcal{F}_t \vee \sigma(L_s, s \leq t).$$

This insider information is considered in BLANCHET-SCALLIET, HILLAIRET & JIAO [16] and will not be considered in this thesis.

S-investor's information

For public bond investors we distinguish two information structures. The first type of investors observe the firm value and the default in the moment it occurs, but they do not have knowledge on the default threshold L (because it is firm inside information of the management). We call an investor endowed with this information structure an *S-investor*. The S-investor's information structure $\mathbb{G}^S = (\mathcal{G}_t^S)_{t \in [0, T]}$ on the bond market is described by a progressive enlargement of the filtration \mathbb{F} by the random default time τ , i.e.,

$$\mathcal{G}_t^S = \mathcal{F}_t \vee \sigma(H_s, s \leq t),$$

where H is the default indicator process defined by $H_t = \mathbf{1}_{\{t \geq \tau\}}$. S-investors are uncertain about the firm's true distance to default although they have complete information about the firm value. This uncertainty is due to lacking knowledge on the threshold level. Thus, default arrives as a full surprise.

D-investor's information

The second type of investors is called a *D-investor*. The information structure of a D-investor is similar to the information structure of an S-investor in that both investors do not have any knowledge about the default barrier except that they observe the occurrence and timing of default. The difference is that the asset process X is not completely observable but only at discrete dates denoted by T_j , $j = 0, \dots, J-1$, where $0 = T_0 < T_1 < \dots < T_{J-1} < T_J = T$ for $J \in \mathbb{N}$. This is a realistic assumption, since investors usually observe the asset value at the times of corporate news release. The partial information on the asset process is described by a sub-filtration $\mathbb{F}^D = (\mathcal{F}_t^D)_{t \in [0, T]}$ of \mathbb{F} , where

$$\mathcal{F}_t^D = \begin{cases} \mathcal{F}_0, & \text{if } t < T_1, \\ \sigma(X_{T_1}, \dots, X_{T_i}) & \text{if } T_i \leq t < T_{i+1}, i \in \{1, \dots, J-1\}. \end{cases}$$

The D-investor's information structure $\mathbb{G}^D = (\mathcal{G}_t^D)_{t \in [0, T]}$ can be described by a progressive enlargement of the filtration \mathbb{F}^D by the random default time τ , i.e.,

$$\mathcal{G}_t^D = \mathcal{F}_t^D \vee \sigma(H_s, s \leq t).$$

Assumption 3.2.1. Every adjustment time t_i , $i = 0, \dots, n-1$, of the default barrier coincides with one of the information dates T_j , $j = 0, \dots, J-1$, i.e., $\{t_0, \dots, t_{n-1}\} \subseteq \{T_0, \dots, T_{J-1}\}$.

3.3 Conditional survival probability

In this section we derive explicit formulas for the conditional survival probability given complete information on the firm's assets and incomplete information on the firm's assets, respectively. Based on these formulas a direct application is the valuation of bond prices. We begin with reviewing classical results of the running minimum of a geometric Brownian motion in Subsection 3.3.1. Then we proceed by deriving explicit formulas for the conditional survival probability given the information of an S-investor and of a D-investor in Subsection 3.3.2 and Subsection 3.3.3, respectively.

3.3.1 Preliminaries

For a geometric Brownian motion $(Y_u)_{u \geq 0}$ given by $Y_u = \exp\{mu + \sigma \widehat{B}_u\}$, where $(\widehat{B}_u)_{u \geq 0}$ is a standard Brownian motion, and its running minimum process $(\widehat{M}_u)_{u \geq 0}$ the joint density function of Y_t and \widehat{M}_t is denoted by $f_t^{\widehat{M}, Y}(\cdot, \cdot)$ and the probability density function of \widehat{M}_t is denoted by $f_t^{\widehat{M}}(\cdot)$. For both density functions there exist explicit expressions given in the following lemma.

Lemma 3.3.1.

1. Given $t > 0$ then the density function of \widehat{M}_t is given by

$$f_t^{\widehat{M}}(w) = \frac{1}{\sigma w \sqrt{t}} \varphi\left(\frac{mt - \frac{\ln w}{\sigma}}{\sqrt{t}}\right) + \frac{e^{2m \frac{\ln w}{\sigma}}}{\sigma w \sqrt{t}} \varphi\left(\frac{mt + \frac{\ln w}{\sigma}}{\sqrt{t}}\right) + \frac{2m}{w\sigma} e^{2m \frac{\ln w}{\sigma}} \Phi\left(\frac{mt + \frac{\ln w}{\sigma}}{\sqrt{t}}\right) \quad (3.3)$$

for $w \in (0, 1]$ and zero otherwise, where Φ and φ denote the cumulative distribution function and the probability density function of the standard normal distribution, respectively.

2. Given $t > 0$ then the joint probability density function of Y_t and its running minimum \widehat{M}_t is given by

$$f_t^{\widehat{M}, Y}(u, v) = \frac{2v^{m/\sigma^2 - 1} \ln(v/u^2)}{\sigma^3 \sqrt{2\pi} t^{3/2} u} e^{-\frac{m^2 t}{2\sigma^2}} e^{-\frac{\ln^2(v/u^2)}{2\sigma^2 t}} \quad (3.4)$$

for $u \in (0, 1]$, $u \leq v$ and zero otherwise.

Proof. The above formulas are a corollary to the results given in HARRISON [39, Ch. 1]. \square

We also introduce the complementary distribution function Ψ of the running minimum \widehat{M}_t , i.e., $\Psi(t, u) = \mathbb{P}(\widehat{M}_t > u)$.

Lemma 3.3.2. *The complementary distribution function Ψ is given by*

$$\Psi(t, u) = \Phi\left(\frac{-\ln(u) + mt}{\sigma\sqrt{t}}\right) - \exp\left\{\frac{2m}{\sigma^2}\ln(u)\right\}\Phi\left(\frac{\ln(u) + mt}{\sigma\sqrt{t}}\right), \quad \text{for } t > 0, u \leq 1, \quad (3.5)$$

$\Psi(t, u) = 0$ for $t \geq 0, u > 1$ and $\Psi(0, u) = 1$ for $u \leq 1$.

Proof. The proof is given in Section 3.5.1. □

3.3.2 S-investor's information case

The S-investor's information structure is given by a progressive enlargement of the filtration \mathbb{F} by the random time τ , where \mathbb{F} is the filtration generated by the asset process. The next theorem shows that the conditional survival probability given the information of an S-investor can be formulated in terms of \mathbb{F} -conditional survival probabilities for which we derive explicit formulas.

Theorem 3.3.3. *For the conditional survival probability given the information of an S-investor it holds*

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t)}{\mathbb{P}(\tau > t | \mathcal{F}_t)}, \quad \text{for } t < T, \quad (3.6)$$

where the \mathbb{F} -conditional survival probabilities are given by the following formulas for $n \geq 2$:

1. For $t \in [t_{i-1}, t_i), i = 1, \dots, n - 1$, we have

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t) = & \int_{[0,1] \times [u, \infty)} \dots \int_{[0,1] \times [u, \infty)} f_{t_i-t}^{\widehat{M}^t, Y^t}(u_i, v_i) \prod_{j=i+1}^{n-1} f_{t_j-t_{j-1}}^{\widehat{M}^{t_{j-1}}, Y^{t_{j-1}}}(u_j, v_j) \\ & \int_0^1 F_{L^1, \dots, L^n}(M_{t_1}, M_{[t_1, t_2)}, \dots, M_{[t_{i-2}, t_{i-1})}, \min(M_{[t_{i-1}, t)}, u_i X_t), \\ & \quad u_{i+1} v_i X_t, \dots, u_{n-1} v_{n-2} \dots v_i X_t, w v_{n-1} \dots v_i X_t) \\ & f_{T-t_{n-1}}^{\widehat{M}^{t_{n-1}}}(w) dw dv_{n-1} du_{n-1} \dots dv_i du_i, \end{aligned} \quad (3.7)$$

For $t \in [t_{n-1}, T)$ it holds

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t) = & \int_0^1 F_{L^1, \dots, L^n}(M_{t_1}, M_{[t_1, t_2)}, \dots, M_{[t_{n-2}, t_{n-1})}, \min(M_{[t_{n-1}, t)}, w X_t)) \\ & f_{T-t}^{\widehat{M}^t}(w) dw. \end{aligned} \quad (3.8)$$

2. For $t \in [t_{i-1}, t_i)$, $i = 1, \dots, n$, it holds

$$\mathbb{P}(\tau > t | \mathcal{F}_t) = F_{L^1, \dots, L^i}(M_{t_1}, M_{[t_1, t_2)}, \dots, M_{[t_{i-2}, t_{i-1})}, M_{[t_{i-1}, t)}).$$

Proof. We obtain Eq. (3.6) by using classical results of progressive enlargement (see JEANBLANC, YOR & CHESNEY [48, Sec. 7.3.3]). For the convenience of the reader the proof of (3.6) is reproduced in Section 3.5.2. For the sake of simpler notation the proof of the \mathbb{F} -conditional survival probabilities is only given for $n = 2$, i.e., the threshold is L^1 in the interval $[t_0, t_1)$ and L^2 in the interval $[t_1, T)$. The proof for $n > 2$ is along the same line and skipped.

Eq. (3.2) yields that for $s > t$ the firm value X_s can be expressed by

$$\begin{aligned} X_s &= X_t \exp\{m(s-t) + \sigma(B_s - B_t)\} = X_t \exp\{m(s-t) + \sigma(B_{s-t+t} - B_t)\} \\ &= X_t \exp\{m(s-t) + \sigma \widehat{B}_{s-t}^t\} = X_t Y_{s-t}^t, \end{aligned}$$

where for any $t \in [t_0, T)$ the process $(\widehat{B}_u^t)_{u \geq 0}$ given by $\widehat{B}_u^t = B_{u+t} - B_t$ is a Brownian motion starting at zero and independent of \mathcal{F}_t and the process $(Y_u^t)_{u \geq 0}$ given by $Y_u^t = \exp\{mu + \sigma \widehat{B}_u^t\}$ is independent of \mathcal{F}_t and it holds

$$Y_u^t = \frac{X_{t+u}}{X_t} \stackrel{d}{=} X_u.$$

1. For the proof of (3.7) let $t \in [t_0, t_1)$ be fixed, then we can describe the event that no default occurs until the maturity time T by

$$\begin{aligned} \{\tau > T\} &= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, T)}\} = \{L^1 < \inf_{s < t_1} X_s\} \cap \{L^2 < \inf_{t_1 \leq s < T} X_s\} \\ &= \{L^1 < \inf_{s < t} X_s\} \cap \{L^1 < \inf_{t \leq s < t_1} X_s\} \cap \{L^2 < \inf_{t_1 \leq s < T} X_s\} \\ &= \{L^1 < \inf_{s < t} X_s\} \cap \{L^1 < \inf_{t \leq s < t_1} X_t Y_{s-t}^t\} \cap \{L^2 < \inf_{t_1 \leq s < T} X_{t_1} Y_{s-t_1}^{t_1}\}, \end{aligned}$$

For any $v \in [t_0, T)$ we denote by $(\widehat{M}_u^v)_{u \geq 0}$ the running minimum process of $(Y_u^v)_{u \geq 0}$ given by

$$\widehat{M}_u^v = \inf_{s < u} Y_s^v.$$

Explicit expressions for the joint density function $f_t^{\widehat{M}^v, Y^v}$ of Y_t^v and \widehat{M}_t^v and for the probability density function $f_t^{\widehat{M}^v}$ of \widehat{M}_t^v are provided in Lemma 3.3.1.

It holds

$$\begin{aligned} \{\tau > T\} &= \{L^1 < M_t\} \cap \{L^1 < \widehat{M}_{t_1-t}^t X_t\} \cap \{L^2 < \widehat{M}_{T-t_1}^{t_1} X_{t_1}\} \\ &= \{L^1 < M_t\} \cap \{L^1 < \widehat{M}_{t_1-t}^t X_t\} \cap \{L^2 < \widehat{M}_{T-t_1}^{t_1} Y_{t_1-t}^t X_t\}. \end{aligned}$$

Based on this representation the conditional survival probability until the maturity time T given the information \mathcal{F}_t can be written as

$$\begin{aligned}
\mathbb{P}(\tau > T | \mathcal{F}_t) &= \mathbb{P}(L^1 < M_t, L^1 < \widehat{M}_{t_1-t}^t X_t, L^2 < \widehat{M}_{T-t_1}^{t_1} Y_{t_1-t}^t X_t | \mathcal{F}_t) \\
&= \int_0^1 \int_u^\infty \mathbb{P}(L^1 < M_t, L^1 < u X_t, L^2 < \widehat{M}_{T-t_1}^{t_1} v X_t | \mathcal{F}_t) f_{t_1-t}^{\widehat{M}^t, Y^t}(u, v) dv du \\
&= \int_0^1 \int_u^\infty \int_0^1 \mathbb{P}(L^1 < \min(M_t, u X_t), L^2 < w v X_t | \mathcal{F}_t) f_{T-t_1}^{\widehat{M}^{t_1}}(w) \\
&\quad f_{t_1-t}^{\widehat{M}^t, Y^t}(u, v) dw dv du \\
&= \int_0^1 \int_u^\infty \int_0^1 F_{L^1, L^2}(\min(M_t, u X_t), w v X_t) f_{T-t_1}^{\widehat{M}^{t_1}}(w) \\
&\quad f_{t_1-t}^{\widehat{M}^t, Y^t}(u, v) dw dv du
\end{aligned}$$

We have exploited in the last equation the independence of (L^1, L^2) from \mathcal{F}_T . For the proof of (3.8) let $t \in [t_1, T)$ be fixed, then we can describe the event that no default occurs until the maturity time T by

$$\begin{aligned}
\{\tau > T\} &= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, T]}\} = \{L^1 < M_{t_1}\} \cap \{L^2 < \inf_{t_1 \leq s < T} X_s\} \\
&= \{L^1 < M_{t_1}\} \cap \{L^2 < \inf_{t_1 \leq s < t} X_s\} \cap \{L^2 < \inf_{t \leq s < T} X_s\} \\
&= \{L^1 < M_{t_1}\} \cap \{L^2 < \inf_{t_1 \leq s < t} X_s\} \cap \{L^2 < \inf_{t \leq s < T} X_t Y_{s-t}^t\} \\
&= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, t]}\} \cap \{L^2 < \widehat{M}_{T-t}^t X_t\}.
\end{aligned}$$

Based on this representation the conditional survival probability until the maturity time T given the information \mathcal{F}_t can be calculated by

$$\begin{aligned}
\mathbb{P}(\tau > T | \mathcal{F}_t) &= \mathbb{P}(L^1 < M_{t_1}, L^2 < M_{[t_1, t]}, L^2 < \widehat{M}_{T-t}^t X_t | \mathcal{F}_t) \\
&= \int_0^1 \mathbb{P}(L^1 < M_{t_1}, L^2 < M_{[t_1, t]}, L^2 < w X_t | \mathcal{F}_t) f_{T-t}^{\widehat{M}^t}(w) dw \\
&= \int_0^1 F_{L^1, L^2}(M_{t_1}, \min(M_{[t_1, t]}, w X_t)) f_{T-t}^{\widehat{M}^t}(w) dw.
\end{aligned}$$

2. The conditional survival probability until time $t \in [t_0, t_1)$ given the information \mathcal{F}_t is obtained by

$$\mathbb{P}(\tau > t | \mathcal{F}_t) = \mathbb{P}(L^1 < M_t | \mathcal{F}_t) = F_{L^1}(M_t).$$

Finally, for $t \in [t_1, T)$ it holds

$$\mathbb{P}(\tau > t | \mathcal{F}_t) = \mathbb{P}(L^1 < M_{t_1}, L^2 < M_{[t_1, t]} | \mathcal{F}_t) = F_{L^1, L^2}(M_{t_1}, M_{[t_1, t]}).$$

□

Remark 3.3.4. For $n = 1$ we obtain as a special case the model proposed by GIESECKE & GOLDBERG [36], where the default barrier is constant but random, i.e., $L_t = L^1$. The conditional survival probability is given by

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t)}{\mathbb{P}(\tau > t | \mathcal{F}_t)} = \mathbf{1}_{\{\tau > t\}} \frac{\int_0^1 F_{L^1}(\min(M_t, wX_t)) f_{T-t}^{\widehat{M}^t}(w) dw}{F_{L^1}(M_t)} \quad (3.9)$$

for $t \in [0, T)$.

Remark 3.3.5. Let $n = 2$ and assume that L^1 and L^2 are independent with probability distribution functions F_{L^1} and F_{L^2} , respectively. Then the conditional survival probability for $t \in [t_0, t_1)$ is given by

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \frac{1}{F_{L^1}(M_t)} \int_0^1 \int_u^\infty \int_0^1 F_{L^1}(\min(M_t, uX_t)) F_{L^2}(wvX_t) f_{T-t_1}^{\widehat{M}^{t_1}}(w) f_{t_1-t}^{\widehat{M}^t, Y^t}(u, v) dw dv du$$

on the no default set $\{\tau > t\}$. For $t \in [t_1, T)$ calculating the conditional survival probability reduces to the case of a constant but random barrier. We have

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t) &= \int_0^1 F_{L^1, L^2}(M_{t_1}, \min(M_{[t_1, t]}, wX_t)) f_{T-t}^{\widehat{M}^t}(w) dw \\ &= F_{L^1}(M_{t_1}) \int_0^1 F_{L^2}(\min(M_{[t_1, t]}, wX_t)) f_{T-t}^{\widehat{M}^t}(w) dw \end{aligned}$$

and

$$\mathbb{P}(\tau > t | \mathcal{F}_t) = F_{L^1, L^2}(M_{t_1}, M_{[t_1, t]}) = F_{L^1}(M_{t_1}) F_{L^2}(M_{[t_1, t]})$$

yielding the following simplified formula for the conditional survival probability

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t)}{\mathbb{P}(\tau > t | \mathcal{F}_t)} = \mathbf{1}_{\{\tau > t\}} \frac{\int_0^1 F_{L^2}(\min(M_{[t_1, t]}, wX_t)) f_{T-t}^{\widehat{M}^t}(w) dw}{F_{L^2}(M_{[t_1, t]})}.$$

A direct application of the conditional survival probability is the pricing of credit derivatives such as defaultable bonds. For example, let us consider a zero-coupon bond that matures at T and has zero recovery, i.e., the defaultable bond pays 1 at T if there was no default by T and zero otherwise. Assuming that the pricing probability is \mathbb{P} , then the price C_t of such a product is given by

$$C_t = e^{-r(T-t)} \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] = e^{-r(T-t)} \mathbb{P}(\tau > T | \mathcal{G}_t^S) = e^{-r(T-t)} \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t)}{\mathbb{P}(\tau > t | \mathcal{F}_t)},$$

where $r \geq 0$ is a discount factor. An important quantity in the credit risk analysis is the credit yield spread S_t on a zero-coupon bond issued by a firm. It is the difference between the yield at time t on a credit risky and a credit risk-free zero-coupon bond, both maturing at T . Thus, the credit spread S_t is given by

$$S_t = -\frac{1}{T-t} \ln\{\mathbb{P}(\tau > T | \mathcal{G}_t^S)\}.$$

3.3.3 D-investor's information case

The D-investor's information structure is given by a progressive enlargement of the filtration \mathbb{F}^D by the random time τ , where \mathbb{F}^D is a sub-filtration of \mathbb{F} . The next theorem shows that the conditional survival probability given the information of a D-investor can be formulated in terms of \mathbb{F}^D -conditional survival probabilities for which explicit formulas are derived. Note, the D-investor obtains information about the asset value only at discrete times which include the adjustment times t_k , $k \in \{0, \dots, n-1\}$, of the default barrier. We denote the information dates between two adjustment times t_k and t_{k+1} by T_i^k , $k = 0, \dots, n-1$, $i = 0, \dots, J_k - 1$, where $T_0^k := t_k$, $T_{i-1}^k < T_i^k$, $T_{J_k-1}^k < t_{k+1}$, $T_{J_k}^k := t_{k+1}$ and $\sum_{k=0}^{n-1} J_k = J$. Further, we introduce

$$K_j^k(\ell) = \mathbb{P}(M_{[T_{j-1}^k, T_j^k]} > \ell | X_{T_{j-1}^k}, X_{T_j^k}) \quad \text{and} \quad K^{k,i}(\ell) = \prod_{j=1}^i K_j^k(\ell)$$

for $k = 0, \dots, n-1$, $j = 1, \dots, J_k$.

Lemma 3.3.6. *The conditional probability $K_j^k(\ell)$, $k = 0, \dots, n-1$, $j = 1, \dots, J_k$, is given by*

$$K_j^k(\ell) = 1 - \exp \left\{ \frac{-2}{\sigma^2(T_j^k - T_{j-1}^k)} \ln \left(\frac{\ell}{X_{T_{j-1}^k}} \right) \ln \left(\frac{\ell}{X_{T_j^k}} \right) \right\},$$

for $\ell < \min(X_{T_{j-1}^k}, X_{T_j^k})$ and $K_j^k(\ell) = 0$ otherwise.

Proof. The proof is given in Section 3.5.3. □

Theorem 3.3.7. *For the conditional survival probability given the information of a D-investor it holds*

$$\mathbb{P}(\tau > T | \mathcal{G}_t^D) = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t^D)}{\mathbb{P}(\tau > t | \mathcal{F}_t^D)} \quad \text{for } t < T, \quad (3.10)$$

where the \mathbb{F}^D -conditional survival probabilities are given by the following formulas for $n \geq 2$:

1. For $t \in [T_i^k, T_{i+1}^k)$, $k = 0, \dots, n-2$, $i = 0, \dots, J_k - 1$, it holds

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t^D) &= \int_0^1 \int_0^\infty \dots \int_0^1 \int_0^\infty f_{T_0^{k+1}-T_i^k}^{\widehat{M}^{T_i^k}, Y^{T_i^k}}(u_{k+1}, v_{k+1}) \prod_{j=k+2}^{n-1} f_{T_0^j-T_0^{j-1}}^{\widehat{M}^{T_0^{j-1}}, Y^{T_0^{j-1}}}(u_j, v_j) \\ &\quad \int_0^1 f_{T-T_0^{n-1}}^{\widehat{M}^{T_0^{n-1}}}(w) \int_0^\infty \int_0^\infty \dots \int_0^\infty \int_0^{u_{k+1}X_{T_i^k}} \int_0^{u_{k+2}v_{k+1}X_{T_i^k}} \dots \\ &\quad \int_0^{u_{n-1}v_{n-2}\dots v_{k+1}X_{T_i^k}} \int_0^{wv_{n-1}\dots v_{k+1}X_{T_i^k}} \prod_{j=0}^{k-1} K^{j, J_j}(\ell^{j+1}) K^{k, i}(\ell^{k+1}) \\ &\quad f_{L^1, \dots, L^n}(\ell^1, \dots, \ell^n) d\ell^n \dots d\ell^1 dw dv_{n-1} du_{n-1} \dots dv_{k+1} du_{k+1}. \end{aligned} \quad (3.A)$$

For $t \in [T_i^{n-1}, T_{i+1}^{n-1})$, $i = 0, \dots, J_{n-1} - 1$, it holds

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t^D) &= \int_0^\infty \dots \int_0^\infty K^{n-1, i}(\ell^n) \prod_{j=0}^{n-2} K^{j, J_j}(\ell^{j+1}) \Psi\left(T - T_i^{n-1}, \frac{\ell^n}{X_{T_i^{n-1}}}\right) \\ &\quad f_{L^1, \dots, L^n}(\ell^1, \dots, \ell^n) d\ell^n \dots d\ell^1. \end{aligned} \quad (3.B)$$

2. For $t \in [T_i^k, T_{i+1}^k)$, $k = 0, \dots, n-1$, $i = 0, \dots, J_k - 1$, it holds

$$\begin{aligned} \mathbb{P}(\tau > t | \mathcal{F}_t^D) &= \int_0^\infty \dots \int_0^\infty \prod_{j=0}^{k-1} K^{j, J_j}(\ell^{j+1}) K^{k, i}(\ell^{k+1}) \Psi\left(t - T_i^k, \frac{\ell^{k+1}}{X_{T_i^k}}\right) \\ &\quad f_{L^1, \dots, L^{k+1}}(\ell^1, \dots, \ell^{k+1}) d\ell^{k+1} \dots d\ell^1. \end{aligned}$$

Proof. The proof is presented in Section 3.5.4. \square

Remark 3.3.8. For the special case of a constant but random default barrier, i.e., $n = 1$ and $L_t = L^1$, the \mathbb{F}^D -conditional survival probabilities are given by the following formulas: Let T_j , $j = 0, \dots, J-1$, where $0 = T_0 < T_1 < \dots < T_{J-1} < T_J = T$ for $J \in \mathbb{N}$, denote the times where the D-investor obtains information about the firm's asset value. For $t \in [T_i, T_{i+1})$, $i = 0, \dots, J-1$, it holds

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{G}_t^D) &= \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t^D)}{\mathbb{P}(\tau > t | \mathcal{F}_t^D)} \\ &= \mathbf{1}_{\{\tau > t\}} \frac{\int_0^1 \int_0^{X_{T_i} w} \prod_{j=1}^i K_j(\ell^1) f_{L^1}(\ell^1) f_{T-T_i}^{\widehat{M}^{T_i}}(w) d\ell^1 dw}{\mathbb{P}(\tau > t | \mathcal{F}_t^D) = \int_0^1 \int_0^{X_{T_i} w} \prod_{j=1}^i K_j(\ell^1) f_{L^1}(\ell^1) f_{t-T_i}^{\widehat{M}^{T_i}}(w) d\ell^1 dw}, \end{aligned} \quad (3.11)$$

where

$$K_j(\ell) = \mathbb{P}(M_{[T_{j-1}, T_j]} > \ell | X_{T_{j-1}}, X_{T_j}).$$

Remark 3.3.9. Let C_t be the price of a zero-coupon bond that matures at T and has zero recovery. Assuming that the pricing probability is \mathbb{P} , then C_t is given by

$$C_t = e^{-r(T-t)} \mathbb{E}[\mathbb{1}_{\{\tau > T\}} | \mathcal{G}_t^D] = e^{-r(T-t)} \mathbb{P}(\tau > T | \mathcal{G}_t^D) = e^{-r(T-t)} \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t^D)}{\mathbb{P}(\tau > t | \mathcal{F}_t^D)},$$

where $r \geq 0$ is a discount factor. The credit spread S_t is given by

$$S_t = -\frac{1}{T-t} \ln\{\mathbb{P}(\tau > T | \mathcal{G}_t^D)\}.$$

3.4 Numerical examples

In this section we implement the formulas derived in the previous section to quantify numerically the impact of asymmetric information on the estimations of the default probabilities. We compare the conditional survival probability given the information of an S-investor, who has complete information about the firm's assets and incomplete information about the default threshold only, to the conditional survival probability given the information of a D-investor, who has incomplete information about both default threshold and firm's assets. In Subsection 3.4.1 we begin with a special but instructive case, where the default threshold L is taken to be constant as proposed in GIESECKE & GOLDBERG [36]. The level L itself is unknown to investors and modeled by making it random. For this case we illustrate and compare the conditional survival probabilities given the information of an S-investor and a D-investor, respectively. Then, in Subsection 3.4.2, the case of a switching default threshold is considered and the conditional survival probabilities given the information of an S-investor and a D-investor are compared. Unless noted otherwise, the parameters for the firm's asset value process are taken from BLANCHET-SCALLIET, HILLAIRET & JIAO [16], i.e.,

$$X_0 = 1, \quad \mu = 0.05 \quad \text{and} \quad \sigma = 0.8.$$

3.4.1 Case of a constant default threshold

In this section we present a special case, where the default threshold L is taken to be constant, i.e., the firm's management decides on a default threshold and makes no adjustment of that threshold. The level L itself is unknown to S-investors and D-investors and modeled by making it random. The conditional survival probabilities given the information of S-investors and D-investors can be numerically calculated by using the formulas (3.9) and (3.11), respectively, where the integrals are evaluated by

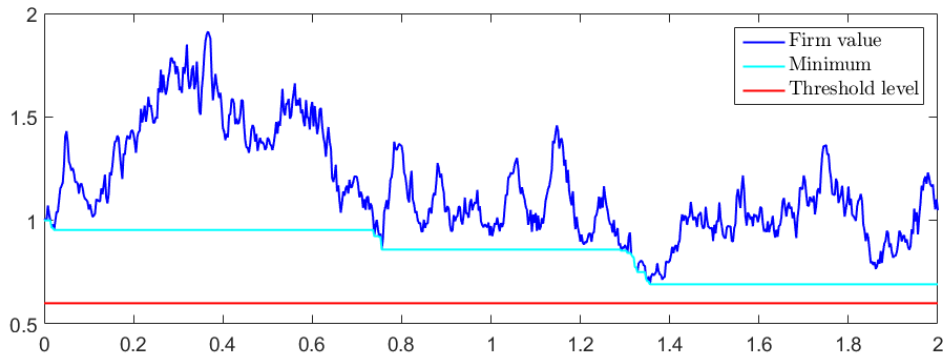
using the Gauss–Kronrod quadrature formula (see MONEGATO [62]).

We present three examples and give numerical comparisons of the conditional default probabilities and credit yield spreads given the information of an S-investor and a D-investor on a given trajectory of the firm’s asset process and a given realization of the default threshold. In all examples S-investors and D-investors assume that the default threshold L follows a standard uniform distribution.

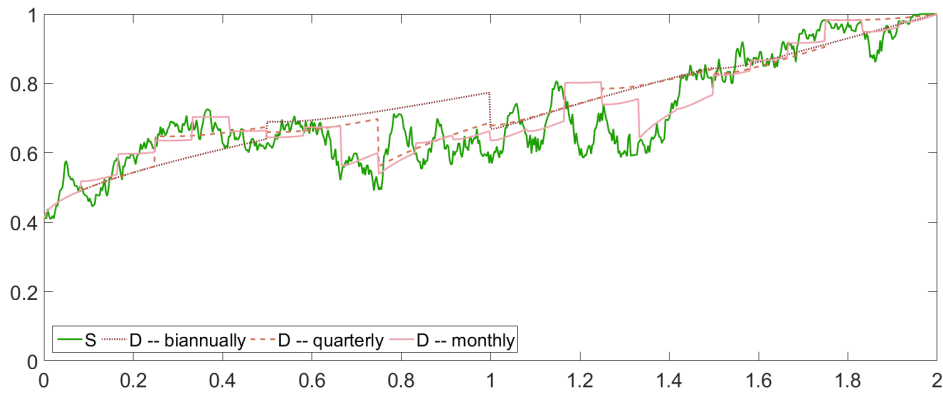
The top panel of Figure 3.2 shows a realized trajectory of the firm value, the associated running minimum and a realization of the default threshold. We observe that the time horizon is $T = 2$ years and the realization of the default threshold is $L(\omega) = 0.6$. No default occurs within the two years. The middle panel of Figure 3.2 illustrates the associated conditional survival probabilities given the information of an S-investor and D-investors. The D-investors obtain information about the firm value twice a year, every quarter and every month, respectively. Since no default occurs within the two years the conditional survival probabilities converge to one. The variation of the conditional survival probability given the information of an S-investor is closely related to the variation of the firm value. This results from the fact that the S-investor perfectly observes the firm value and thus learns about the default threshold, i.e., the S-investor learns that L must lie below the current running minimum of the firm value if default has not yet occurred. Thus, the larger the distance of the firm value to the running minimum the less likely is a default and the S-investor adjusts the estimation of the default probability accordingly. Further, we observe that the more frequently a D-investor obtains information about the firm value the closer are the D-investor’s and S-investor’s estimates of the survival probability. The bottom panel of Figure 3.2 presents the associated credit yield spreads. The credit yield spread given the information of an S-investor is high, if the firm value is close to its running minimum and low otherwise. Especially, the credit spread tends to zero as the time t approaches maturity T . These properties result from the fact that S-investors learn about the default threshold and whenever $X_t > M_t$ the S-investor knows that the firm is not subject to default in the next instance of time since the firm value is continuous and cannot jump to the level M_t . Note that the firm value is far above its running minimum just before maturity and S-investors know that there will be no default by T . Thus, S-investors will not demand a default risk premium and the credit spread is zero. Since D-investors cannot be certain about the firm value just before maturity they demand a risk premium for default risk and credit spreads are non-zero at T . Furthermore, the credit spread at T of a D-investor depends on their last observed firm value. Note that the D-investors who obtain information about the firm value twice a year, every quarter, and every month receive the last information before T at $t = 1.5$, $t = 1.75$, and $t = 1.9167$, respectively. At these time points the firm value is $X_{1.5} = 0.97$, $X_{1.75} = 1.36$, and $X_{1.9167} = 0.94$. Thus, the D-investor who obtains information about the firm value every month demands the highest risk premium at T , whereas the lowest risk premium at T is demanded by the D-investor who obtains information about the firm value every quarter.

The top panel of Figure 3.3 shows a realized trajectory of the firm value, the associated running minimum, and a realization of the default threshold. We observe that the time horizon is $T = 2$ years and the realization of the default threshold is $L(\omega) = 0.6$, which is similar to the previous example. However, this time default occurs in the second year. The middle and bottom panel of Figure 3.3 illustrate the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors, respectively. Similar to the previous example, D-investors obtain information about the firm value twice a year, every quarter, and every month, respectively. We observe that the conditional survival probabilities jump to zero at the time of default. Note that the D-investors who obtain information about the firm value twice a year, every quarter and every month receive the last information before the default τ at $t = 1$, $t = 1.25$, and $t = 1.3333$, respectively. At these time points the firm value is $X_1 = 0.6117$, $X_{1.25} = 0.8655$, and $X_{1.3333} = 0.6046$. Thus, the D-investor who obtains information about the firm value every month demands the highest risk premium just before the default occurs, whereas the lowest risk premium is demanded by the D-investor who obtains information about the firm value every quarter.

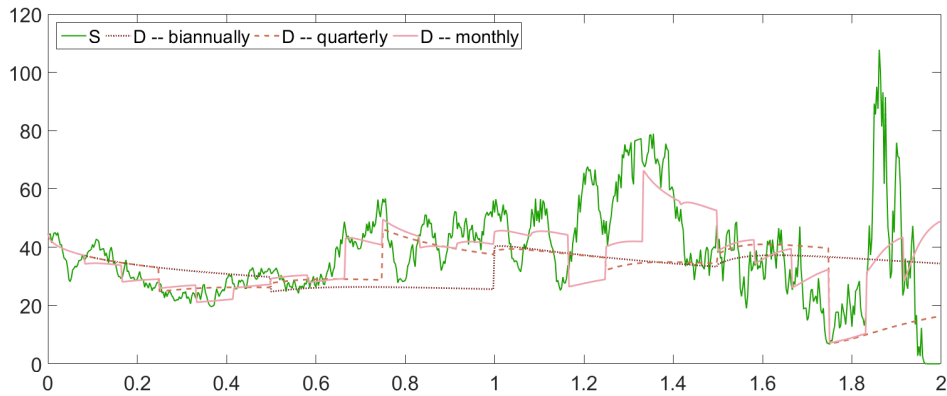
We note that the estimates of the conditional survival probabilities given the information of an S-investor and a D-investor are very close at the D-investor's information dates in the two examples presented above. This is no longer the case if we extend the time horizon. The top panel of Figure 3.4 shows a realized trajectory of the firm value, the associated running minimum, and a realization of the default threshold, where the time horizon is $T = 5$ years. The middle and bottom panel show the associated conditional survival probabilities and credit yield spreads given the information of an S-investor and a D-investor, respectively. The D-investor obtains information about the firm value once in 2.5 years. We observe that the S-investor's estimate of the conditional survival probability and the D-investor's estimate of the conditional survival probability are now visibly different at the information date of the D-investor.



(a) Firm value, running minimum process, and default threshold.

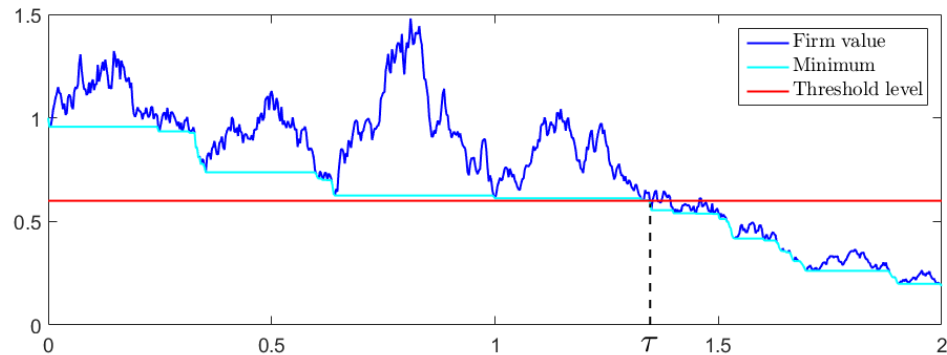


(b) Conditional survival probabilities given the information of an S-investor and D-investors.

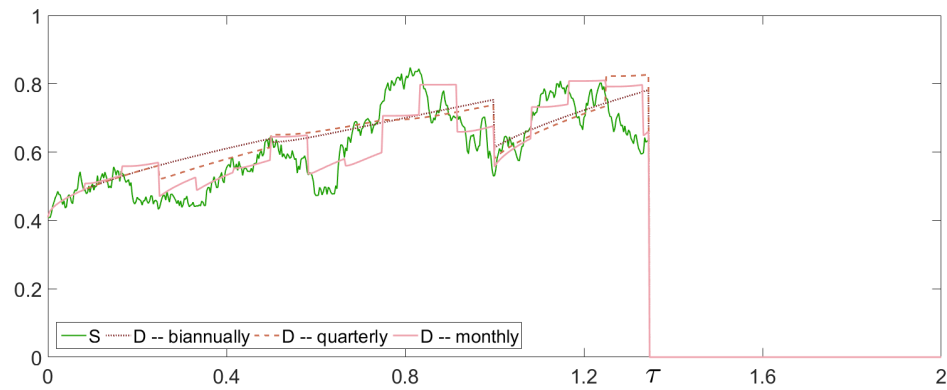


(c) Credit yield spreads given the information of an S-investor and D-investors.

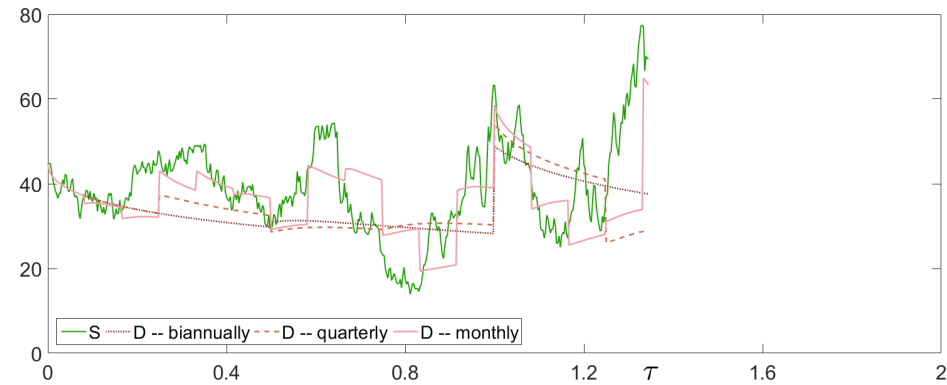
Figure 3.2: Plot of a trajectory of the firm's asset process and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors. The D-investors obtain information about the firm value twice a year, every quarter, and every month, respectively.



(a) Firm value, running minimum process, and default threshold.

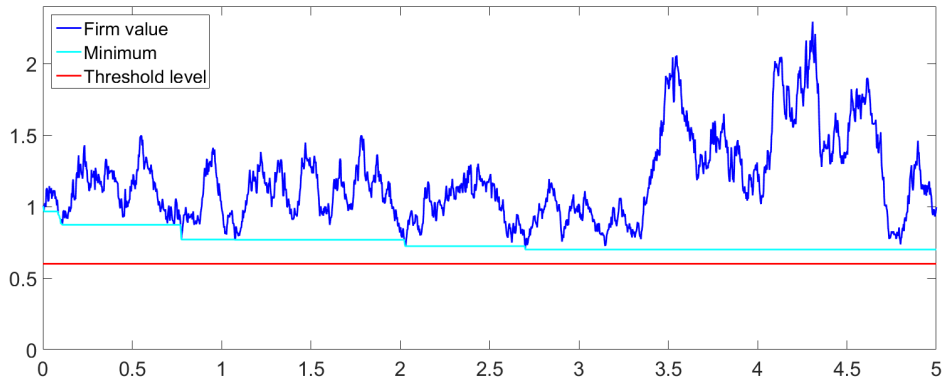


(b) Conditional survival probabilities given the information of an S-investor and D-investors.

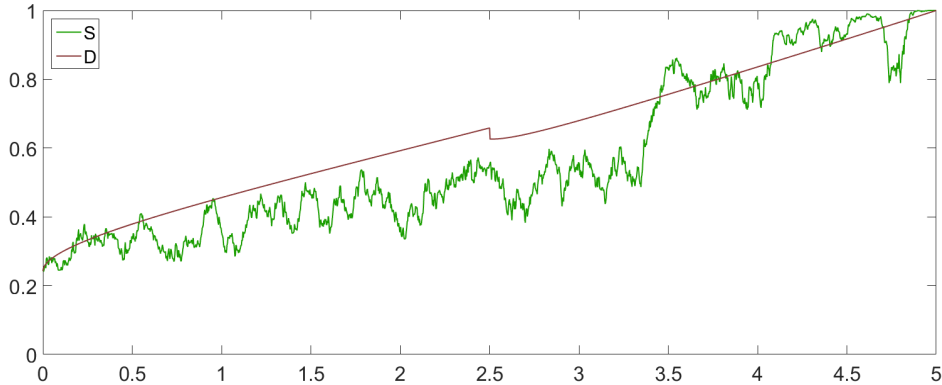


(c) Credit yield spreads given the information of an S-investor and D-investors.

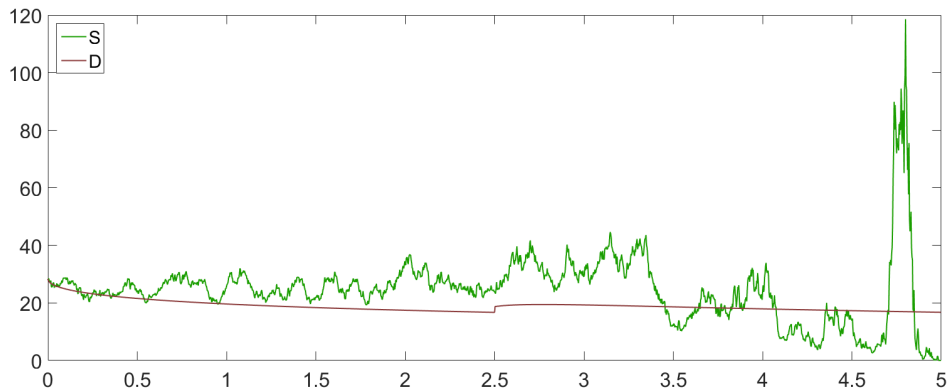
Figure 3.3: Plot of a trajectory of the firm's asset process and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors. The D-investors obtain information about the firm value twice a year, every quarter, and every month, respectively.



(a) Firm value, running minimum process, and default threshold.



(b) Conditional survival probability given the information of an S-investor and a D-investor.



(c) Credit yield spreads given the information of an S-investor and a D-investor.

Figure 3.4: Plot of a trajectory of the firm's asset process and the associated conditional survival probabilities and credit yield spreads given the information of an S-investor and a D-investor. The D-investor obtains information about the firm value once in 2.5 years.

3.4.2 Case of a switching default threshold

In this section we compare the conditional survival probabilities given the information of an S-investor and a D-investor for the case of a switching barrier by numerical examples. We consider a time horizon of $T = 2$ years and we suppose that a firm's management decides at $t_0 = 0$ on a default threshold L^1 . Further, we assume that the management adjusts the default threshold at $t_1 = 1$ from L^1 to L^2 . S-investors and D-investors only have knowledge on the (marginal and joint) laws of $\mathbf{L} = (L^1, L^2)$. The law of $\mathbf{L} = (L^1, L^2)$ is modeled by a copula. The concept of copulas is briefly introduced in this subsection, for a comprehensive introduction to copulas see NELSEN [63].

Definition 3.4.1. Assume that the random vector (U^1, U^2, \dots, U^n) has uniformly distributed marginals, i.e.,

$$U^k \sim U(0, 1), \quad k = 1, \dots, n,$$

then the joint cumulative distribution function of (U^1, U^2, \dots, U^n) is called a *copula* and denoted by C , i.e.,

$$C(u_1, u_2, \dots, u_n) = \mathbb{P}(U^1 \leq u_1, U^2 \leq u_2, \dots, U^n \leq u_n).$$

The next theorem shows the importance of copulas.

Theorem 3.4.2 (Sklar's Theorem). *For any random vector (X^1, X^2, \dots, X^n) with joint cumulative distribution function*

$$F_{X^1, X^2, \dots, X^n}(x_1, x_2, \dots, x_n) = \mathbb{P}(X^1 \leq x_1, X^2 \leq x_2, \dots, X^n \leq x_n)$$

and marginal distribution functions

$$F_{X^k}(x) = \mathbb{P}(X^k \leq x), \quad k = 1, \dots, n,$$

there exists a copula C such that

$$F_{X^1, X^2, \dots, X^n}(x_1, x_2, \dots, x_n) = C(F_{X^1}(x_1), F_{X^2}(x_2), \dots, F_{X^n}(x_n)).$$

If each of the marginal distribution function $F_{X^k}(x)$, $k = 1, 2, \dots, n$, is continuous then C is unique.

Proof. The proof in the case that all the marginal distribution functions are continuous is provided here, for the complete proof see SKLAR [77].

Assuming that $F_{X^k}(\cdot)$, $k = 1, 2, \dots, n$, is continuous and strictly increasing then there exists an inverse function $F_{X^k}^{-1}(\cdot)$ such that

$$F_{X^k}(F_{X^k}^{-1}(u)) = u, \quad \text{for all } 0 \leq u \leq 1.$$

Let $U^k = F_{X^k}(X^k)$, $k = 1, 2, \dots, n$, then it holds

$$\mathbb{P}(U^k \leq u) = \mathbb{P}(F_{X^k}(X^k) \leq u) = \mathbb{P}(X^k \leq F_{X^k}^{-1}(u)) = F_{X^k}(F_{X^k}^{-1}(u)) = u.$$

Thus, U^k , $k = 1, 2, \dots, n$, is uniformly distributed and we have

$$\begin{aligned} F_{X^1, \dots, X^n}(x_1, \dots, x_n) &= \mathbb{P}(X^1 \leq x_1, \dots, X^n \leq x_n) \\ &= \mathbb{P}(F_{X^1}(X^1) \leq F_{X^1}(x_1), \dots, F_{X^n}(X^n) \leq F_{X^n}(x_n)) \\ &= \mathbb{P}(U^1 \leq F_{X^1}(x_1), \dots, U^n \leq F_{X^n}(x_n)) \\ &= C(F_{X^1}(x_1), \dots, F_{X^n}(x_n)), \end{aligned}$$

where C denotes the copula of (U^1, U^2, \dots, U^n) . □

Remark 3.4.3. Sklar's Theorem allows us to describe the joint cumulative distribution function of a random vector (X^1, X^2, \dots, X^n) by its marginal distributions $F_{X^k}(x)$, $k = 1, \dots, n$ and the copula C , which links the marginal distributions together to form the joint distribution.

In the following examples S-investors and D-investors assume that L^1 is beta distributed with parameters $(\alpha, \beta) = (2, 2)$ and L^2 is exponentially distributed with parameter $\lambda = 2/3$. Further, investors assume that the law of \mathbf{L} is given by a Gumbel copula C , i.e.,

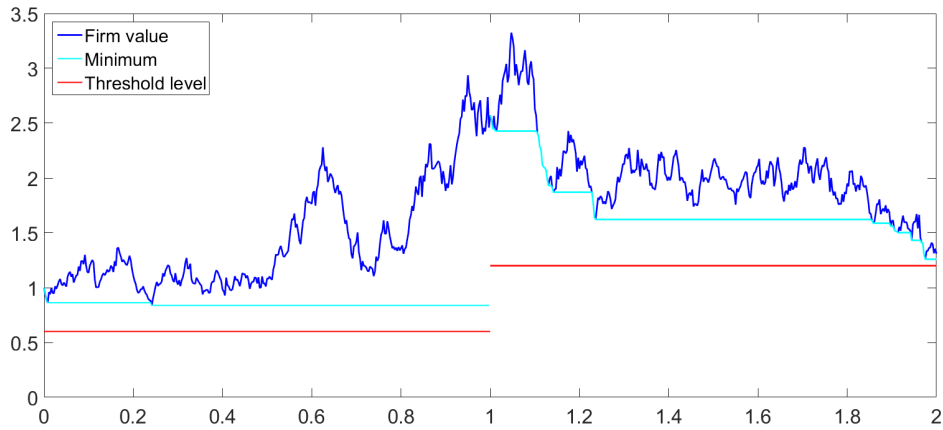
$$C(x_1, x_2) = \exp \left\{ - \left[(-\ln(x_1))^\theta + (-\ln(x_2))^\theta \right]^{\frac{1}{\theta}} \right\},$$

for some $\theta \geq 1$ (see BLUHM & OVERBECK [17]). Thus, the correlation between L^1 and L^2 is modeled by the parameter θ , where $\theta = 1$ corresponds to the case of independent default thresholds.

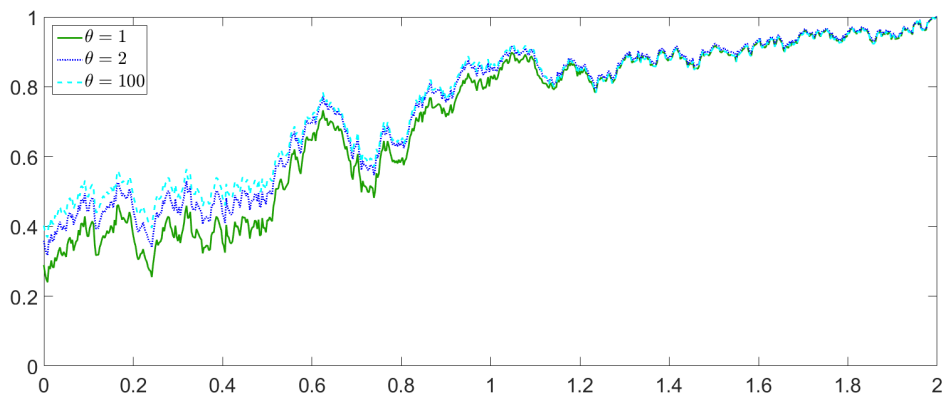
The top panel of Figure 3.5 presents a realized trajectory of the firm's asset process, the switching default threshold and the running minimum of the firm value which is restarted after adjustment of the default threshold. We observe that no default occurs before maturity and the realizations of the default thresholds are given by $L^1(\omega) = 0.6$ and $L^2(\omega) = 1.2$. The middle panel of Figure 3.5 shows the associated conditional survival probability given the information of an S-investor for different values of θ ($\theta = 1, 2, 100$). We observe that the conditional survival probability converges to one, since no default has occurred by maturity. If L^1 and L^2 are independent, i.e., $\theta = 1$, the estimate of the survival probability is smaller compared to the cases of dependent L^1 and L^2 . However, differences in the conditional survival probability between $\theta = 1$, $\theta = 2$, and $\theta = 100$ become smaller with increasing time. The bottom panel of Figure 3.5 illustrates the associated credit yield spread given the information of an S-investor for different values of θ ($\theta = 1, 2, 100$). The credit spread tends to zero as the time t approaches maturity T , since the S-investor has learned about the default threshold and knows that the firm is not subject to default just before T . However, there are

strong fluctuations of the credit spread before it tends to zero resulting from the firm value X_t being close to or attaining its running minimum $M_{[1,t]}$. At times where the firm value is close to or attains its running minimum, the S-investor has to be prepared that the firm might default in the next instance of time and thus, he demands a credit risk premium. If L^1 and L^2 are independent, i.e., $\theta = 1$, the credit yield spread is higher compared to the cases of dependent L^1 and L^2 . Figure 3.6 illustrates the conditional survival probability and the credit spread given the information of a D-investor for the case of independent default thresholds ($\theta = 1$) and different information time points (biannually, quarterly, monthly). We observe that the more frequently a D-investor obtains information about the firm value, the closer are the D-investor's and S-investor's estimates of the survival probability. Similar to the case of a constant default threshold, the credit yield spread given the information of a D-investor is non-zero at maturity, i.e., D-investors demand a risk premium for the default risk. Recall that the D-investors who obtain information about the firm value twice a year, every quarter, and every month receive the last information before T at $t = 1.5$, $t = 1.75$, and $t = 1.9167$, respectively. At these time points the firm value is $X_{1.5} \approx 2.16$, $X_{1.75} \approx 2.16$, and $X_{1.9167} \approx 1.55$. Thus, the D-investor who obtains information about the firm value every month demands the highest risk premium. The case of dependent default thresholds, i.e., $\theta > 1$, leads to qualitatively the same estimates of the conditional survival probability as for $\theta = 1$.

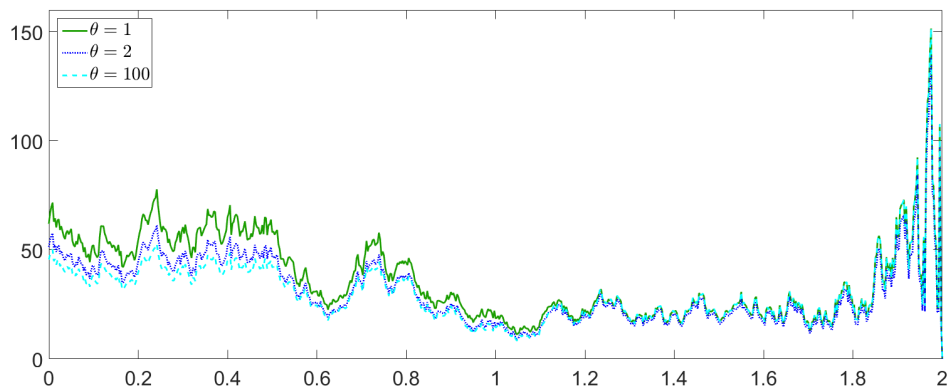
A second example is presented by Figure 3.7. The top panel shows a realized trajectory of the firm's asset process, the switching default threshold and the running minimum of the firm value which is restarted after adjustment of the default threshold. We observe that default occurs in the second year. The middle and bottom panel illustrate the associated conditional survival probability and credit yield spread given the information of an S-investor for different values of θ ($\theta = 1, 2, 100$). We observe that the conditional survival probability jumps to zero at the time of default. Figure 3.8 illustrates the conditional survival probability and credit yield spread given the information of a D-investor for the case of independent default thresholds ($\theta = 1$). The D-investors obtain information about the firm value twice a year, every quarter and every month, respectively. We observe that the conditional survival probabilities jump to zero at the time of default. The case of dependent default thresholds, i.e., $\theta > 1$, leads to qualitatively the same estimates of the conditional survival probability as for $\theta = 1$.



(a) Firm value, running minimum process, and default threshold.

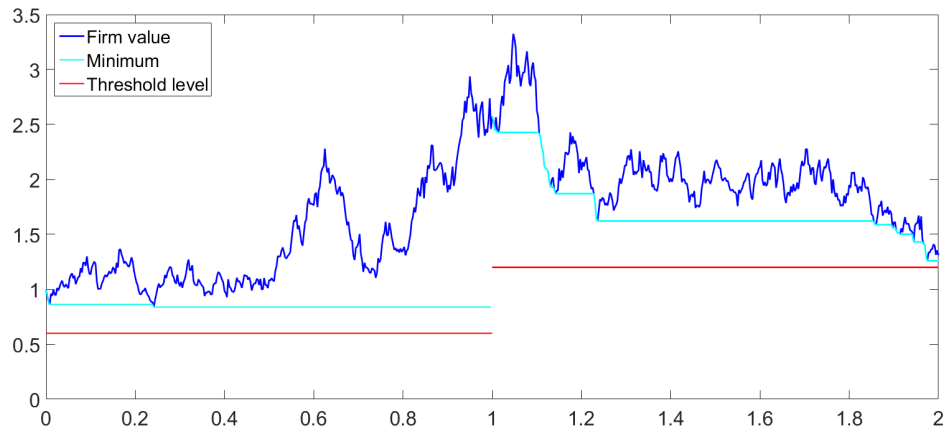


(b) Conditional survival probability given the information of an S-investor.

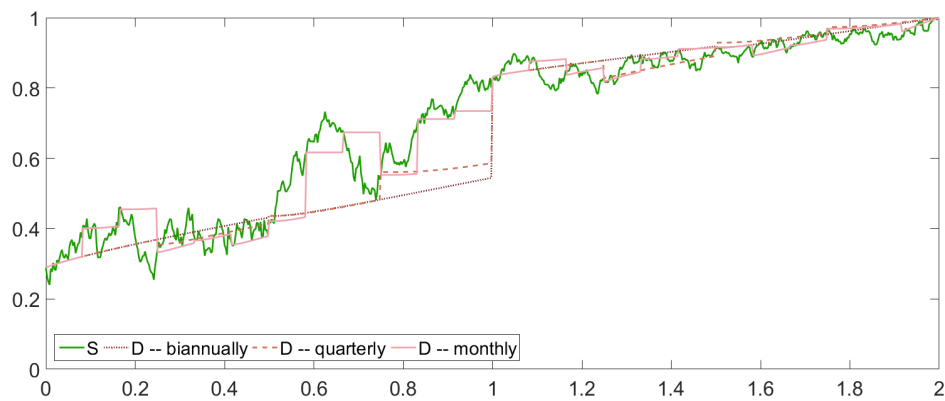


(c) Credit yield spread given the information of an S-investor.

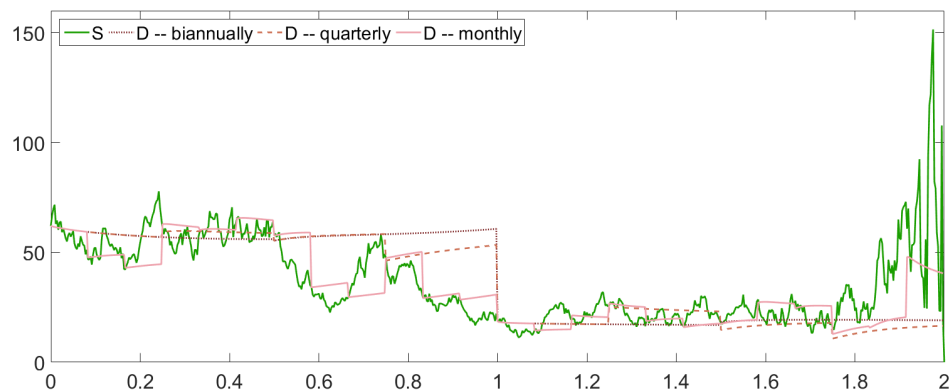
Figure 3.5: Plot of a trajectory of the firm's asset process and the associated conditional survival probability and credit spread given the information of an S-investor for different values of θ .



(a) Firm value, running minimum process, and default threshold.

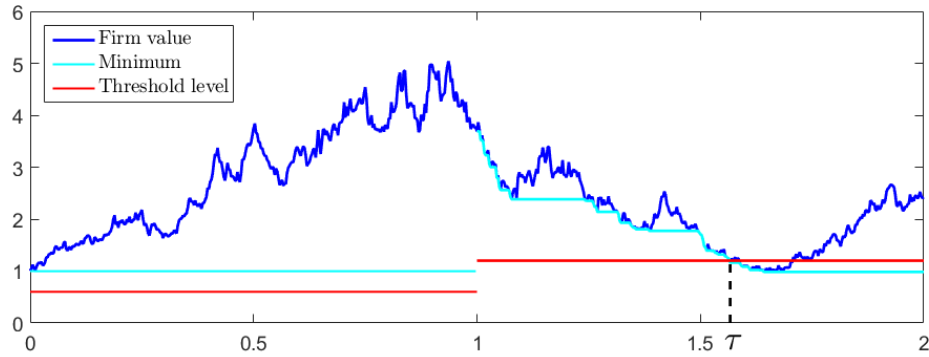


(b) Conditional survival probabilities given the information of an S-investor and D-investors.

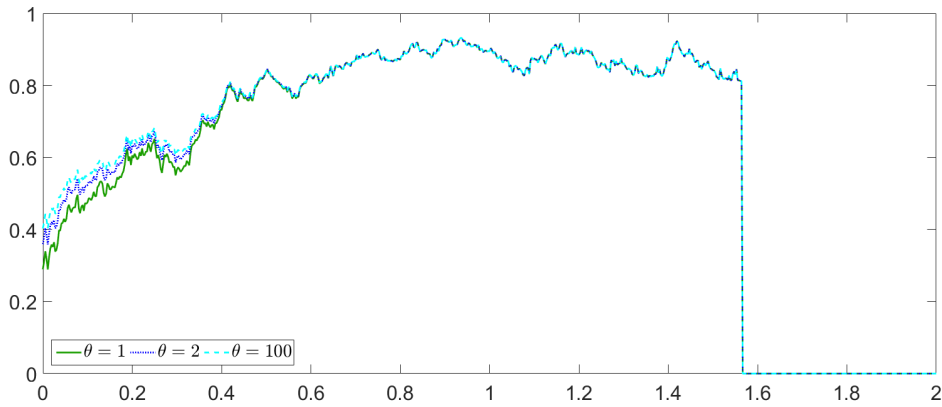


(c) Credit yield spreads given the information of an S-investor and D-investors.

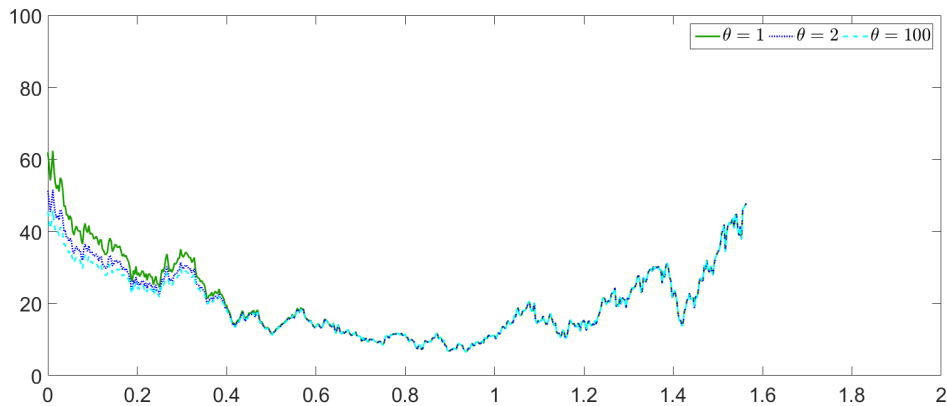
Figure 3.6: Plot of a trajectory of the firm's asset process and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.



(a) Firm value, running minimum process, and default threshold.

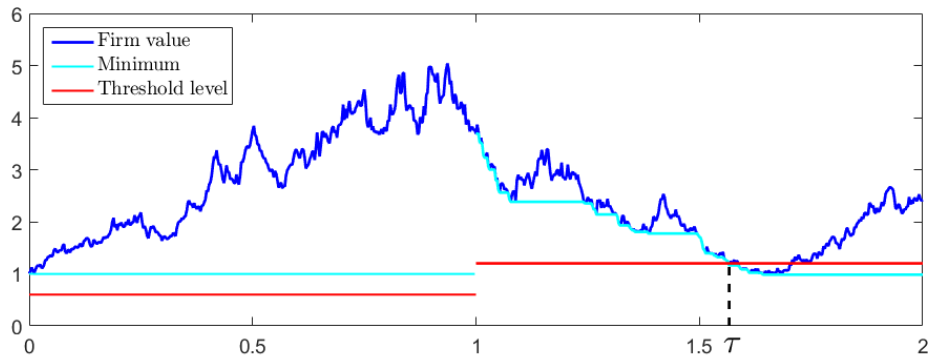


(b) Conditional survival probability given the information of an S-investor.

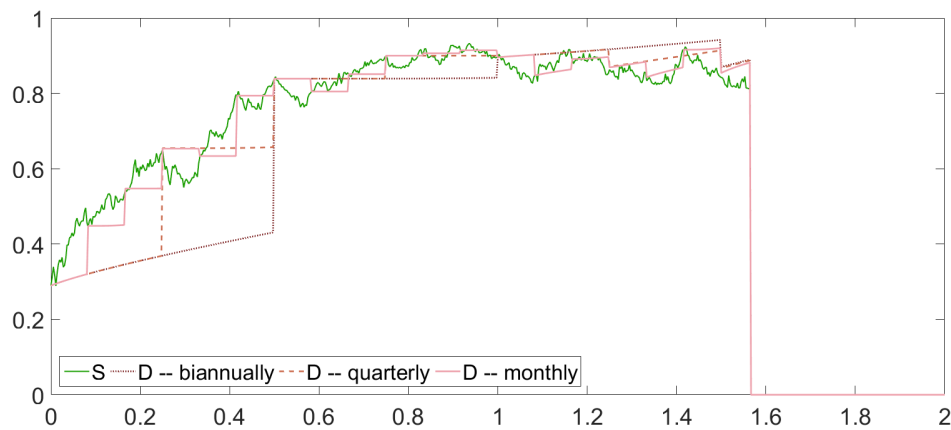


(c) Credit yield spread given the information of an S-investor.

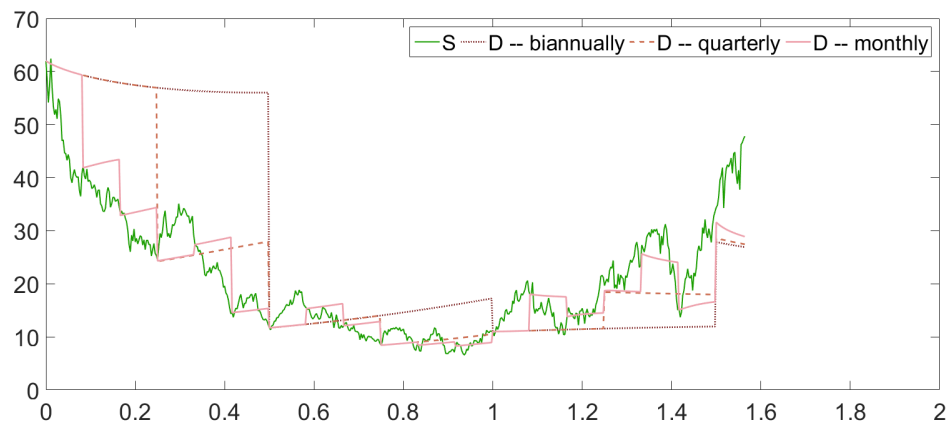
Figure 3.7: Plot of a trajectory of the firm's asset process and the associated conditional survival probability and credit spread given the information of an S-investor for different values of θ .



(a) Firm value, running minimum process, and default threshold.



(b) Conditional survival probabilities given the information of an S-investor and D-investors.



(c) Credit yield spreads given the information of an S-investor and D-investors.

Figure 3.8: Plot of a trajectory of the firm's asset process and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.

3.5 Proofs

3.5.1 Proof of Lemma 3.3.2

The proof is along the same line as in JEANBLANC & VALCHEV [47, Lemma 2]. We have

$$\begin{aligned}\mathbb{P}(\widehat{M}_t > u) &= \mathbb{P}(\widehat{M}_t > u, Y_t > u) = \mathbb{P}(Y_t > u) - \mathbb{P}(\widehat{M}_t \leq u, Y_t > u) \\ &= 1 - \mathbb{P}(Y_t \leq u) - \mathbb{P}(\widehat{M}_t \leq u, Y_t > u).\end{aligned}$$

A classical result from the reflection principle is (see HARRISON [39, Sec. 1.6])

$$\mathbb{P}(\widehat{M}_t \leq u, Y_t > v) = \exp\left\{\frac{2m}{\sigma^2} \ln(u)\right\} \Phi\left(\frac{-\ln(v) + 2\ln(u) + mt}{\sigma\sqrt{t}}\right) \quad (3.12)$$

and from the log-normal distribution of Y_t we obtain

$$\mathbb{P}(Y_t \leq v) = \Phi\left(\frac{\ln(v) - mt}{\sigma\sqrt{t}}\right) \quad (3.13)$$

yielding the desired result

$$\begin{aligned}\mathbb{P}(\widehat{M}_t > u) &= 1 - \Phi\left(\frac{\ln(u) - mt}{\sigma\sqrt{t}}\right) - \exp\left\{\frac{2m}{\sigma^2} \ln(u)\right\} \Phi\left(\frac{\ln(u) + mt}{\sigma\sqrt{t}}\right) \\ &= \Phi\left(\frac{-\ln(u) + mt}{\sigma\sqrt{t}}\right) - \exp\left\{\frac{2m}{\sigma^2} \ln(u)\right\} \Phi\left(\frac{\ln(u) + mt}{\sigma\sqrt{t}}\right).\end{aligned}$$

3.5.2 Proof of Equation (3.6)

The S-investor's information structure $\mathbb{G}^S = (\mathcal{G}_t^S)_{t \in [0, T]}$ is described by a progressive enlargement of the filtration \mathbb{F} by the random default time τ , i.e.,

$$\mathcal{G}_t^S = \mathcal{F}_t \vee \sigma(H_s, s \leq t),$$

where \mathbb{F} is the natural filtration generated by a geometric Brownian motion and H is the default indicator process defined by $H_t = \mathbb{1}_{\{t \geq \tau\}}$. First, we show that on the no default set $\{\tau > t\}$ every element of \mathcal{G}_t^S can be represented by an element of \mathcal{F}_t .

Proposition 3.5.1. *Let \mathcal{K}_t be a collection of sets defined by*

$$\mathcal{K}_t := \{A \in \mathcal{A} \mid \exists B \in \mathcal{F}_t: A \cap \{\tau > t\} = B \cap \{\tau > t\}\}.$$

Then $\mathcal{G}_t^S \subset \mathcal{K}_t$ for every $t \in \mathbb{R}_+$.

Proof. We denote $\mathcal{H}_t = \sigma(\{\tau \leq u\}, u \leq t)$ and $\mathcal{E}_t = \mathcal{H}_t \cup \mathcal{F}_t := \{E = H \cup F \mid H \in \mathcal{H}_t, F \in \mathcal{F}_t\}$. Note that $\mathcal{G}_t^S = \sigma(\mathcal{E}_t)$. We show below that \mathcal{K}_t is a Dynkin system and closed under finite intersections which yields that \mathcal{K}_t is a σ -algebra (see BAUER [6, Theorem 2.3]). Then it is sufficient to show that $\mathcal{E}_t \subset \mathcal{K}_t$ since we have from the properties of σ -algebras that $\mathcal{G}_t^S = \sigma(\mathcal{E}_t) \subset \sigma(\mathcal{K}_t) = \mathcal{K}_t$. Thus, we show that for $E \in \mathcal{E}_t$ there exists an event $B \in \mathcal{F}_t$ such that $E \cap \{\tau > t\} = B \cap \{\tau > t\}$. The set E can be written as $E = H \cup F$ for some $H \in \mathcal{H}_t$ and $F \in \mathcal{F}_t$. Then

$$E \cap \{\tau > t\} = (H \cup F) \cap \{\tau > t\} = (H \cap \{\tau > t\}) \cup (F \cap \{\tau > t\}).$$

It holds $H \cap \{\tau > t\} = \emptyset$ since $H = \{\tau \leq u\}$ for some $u \leq t$. Setting $B = F \in \mathcal{F}_t$ we find

$$E \cap \{\tau > t\} = F \cap \{\tau > t\} = B \cap \{\tau > t\}.$$

Finally it is left to show that \mathcal{K}_t is a Dynkin system closed under finite intersections:

1. $\Omega \in \mathcal{K}_t$, i.e., there exists an event $B \in \mathcal{F}_t$ such that $\Omega \cap \{\tau > t\} = B \cap \{\tau > t\}$. This is obtained by setting $B := \Omega \in \mathcal{F}_t$.

2. For $A_1, A_2 \in \mathcal{K}_t$ and $A_1 \subset A_2$ it holds $A_2 \setminus A_1 \in \mathcal{K}_t$. We have

$$\begin{aligned} (A_2 \setminus A_1) \cap \{\tau > t\} &= (A_2 \cap \{\tau > t\}) \setminus (A_1 \cap \{\tau > t\}) \\ &= (B_2 \cap \{\tau > t\}) \setminus (B_1 \cap \{\tau > t\}) = (B_2 \setminus B_1) \cap \{\tau > t\}, \end{aligned}$$

where $B_1, B_2 \in \mathcal{F}_t$ since $A_1, A_2 \in \mathcal{K}_t$ and $B_2 \setminus B_1 \in \mathcal{F}_t$ since \mathcal{F}_t is a σ -algebra.

3. For a sequence of pairwise disjoint subsets $(A_n)_{n \in \mathbb{N}} \subset \mathcal{K}_t$ we have $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{K}_t$ since

$$\begin{aligned} \bigcup_{n \in \mathbb{N}} A_n \cap \{\tau > t\} &= \bigcup_{n \in \mathbb{N}} (A_n \cap \{\tau > t\}) = \bigcup_{n \in \mathbb{N}} (B_n \cap \{\tau > t\}) \\ &= \left(\bigcup_{n \in \mathbb{N}} B_n \right) \cap \{\tau > t\}, \end{aligned}$$

where $B_n \in \mathcal{F}_t$ for all $n \in \mathbb{N}$ and $\bigcup_{n \in \mathbb{N}} B_n \in \mathcal{F}_t$ since \mathcal{F}_t is a σ -algebra.

From the above results we obtain that \mathcal{K}_t is a Dynkin system. Let $A_1, A_2 \in \mathcal{K}_t$ then

$$\begin{aligned} (A_1 \cap A_2) \cap \{\tau > t\} &= A_1 \cap (A_2 \cap \{\tau > t\}) = A_1 \cap (B_2 \cap \{\tau > t\}) \\ &= (A_1 \cap \{\tau > t\}) \cap B_2 = (B_1 \cap \{\tau > t\}) \cap B_2 \\ &= (B_1 \cap B_2) \cap \{\tau > t\}, \end{aligned}$$

where $B_1, B_2 \in \mathcal{F}_t$ and thus $B_1 \cap B_2 \in \mathcal{F}_t$. It follows that \mathcal{K}_t is closed under finite intersections. □

Using the above representation of the set \mathcal{G}_t^S we can show the following Lemma.

Lemma 3.5.2. *Let Y be an integrable random variable. Then*

$$\mathbf{1}_{\{\tau > t\}} \mathbb{E}[Y | \mathcal{G}_t^S] = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[Y \mathbf{1}_{\{\tau > t\}} | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)}.$$

Proof. We denote $C = \{\tau > t\}$ and let $A \in \mathcal{G}_t^S$. From Proposition 3.5.1 we have that $A \cap C = B \cap C$ for some $B \in \mathcal{F}_t$. For the proof of the lemma we have to verify that

$$\mathbf{1}_C \mathbb{E}[Y | \mathcal{G}_t^S] \mathbb{P}(C | \mathcal{F}_t) = \mathbf{1}_C \mathbb{E}[Y \mathbf{1}_C | \mathcal{F}_t].$$

We obtain for the left-hand side

$$\mathbf{1}_C \mathbb{E}[Y | \mathcal{G}_t^S] \mathbb{P}(C | \mathcal{F}_t) = \mathbb{E}[\mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) | \mathcal{G}_t^S]$$

and for the right-hand side

$$\mathbf{1}_C \mathbb{E}[Y \mathbf{1}_C | \mathcal{F}_t] = \mathbb{E}[\mathbf{1}_C \mathbb{E}[Y \mathbf{1}_C | \mathcal{F}_t] | \mathcal{G}_t^S],$$

where we have used that \mathcal{F}_t and $\mathcal{H}_t = \sigma(\{\tau \leq u\}, u \leq t)$ are sub-filtrations of \mathcal{G}_t^S . Thus, we have to show that

$$\mathbb{E}[\mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) | \mathcal{G}_t^S] = \mathbb{E}[\mathbf{1}_C \mathbb{E}[Y \mathbf{1}_C | \mathcal{F}_t] | \mathcal{G}_t^S],$$

i.e., for any $A \in \mathcal{G}_t^S$ we have to prove that

$$\int_A \mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} = \int_A \mathbf{1}_C \mathbb{E}[Y \mathbf{1}_C | \mathcal{F}_t] d\mathbb{P}.$$

Using that $A \cap C = B \cap C$ for some $B \in \mathcal{F}_t$ we have

$$\int_A \mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} = \int_{A \cap C} Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} = \int_{B \cap C} Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} = \int_B \mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P}.$$

From the definition of the conditional expectation we have

$$\begin{aligned} \int_B \mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} &= \int_B \mathbb{E}[\mathbf{1}_C Y \mathbb{P}(C | \mathcal{F}_t) | \mathcal{F}_t] d\mathbb{P} = \int_B \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] \mathbb{P}(C | \mathcal{F}_t) d\mathbb{P} \\ &= \int_B \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] \mathbb{E}[\mathbf{1}_C | \mathcal{F}_t] d\mathbb{P} = \int_B \mathbb{E}[\mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] \mathbf{1}_C | \mathcal{F}_t] d\mathbb{P}. \end{aligned}$$

In the second equation and in the last equation we have used that $\mathbb{P}(C | \mathcal{F}_t)$ and $\mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t]$ are \mathcal{F}_t -measurable, respectively. Using again the definition of the conditional expectation we obtain

$$\begin{aligned} \int_B \mathbb{E}[\mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] \mathbf{1}_C | \mathcal{F}_t] d\mathbb{P} &= \int_B \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] \mathbf{1}_C d\mathbb{P} = \int_{B \cap C} \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] d\mathbb{P} \\ &= \int_{A \cap C} \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] d\mathbb{P} = \int_A \mathbf{1}_C \mathbb{E}[\mathbf{1}_C Y | \mathcal{F}_t] d\mathbb{P}, \end{aligned}$$

leading to the desired result. \square

Corollary 3.5.3. *The conditional survival probability given the information of an S -investor is obtained by Eq. (3.6), i.e.,*

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{P}(\tau > T | \mathcal{F}_t)}{\mathbb{P}(\tau > t | \mathcal{F}_t)}, \quad \text{for } t < T.$$

Proof. We have

$$\mathbb{P}(\tau > T | \mathcal{G}_t^S) = \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] = \mathbf{1}_{\{\tau > t\}} \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] + \mathbf{1}_{\{\tau \leq t\}} \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S].$$

The second term on the right-hand side vanishes because $\mathbf{1}_{\{\tau \leq t\}}$ is \mathcal{G}_t^S -measurable, leading to

$$\mathbf{1}_{\{\tau \leq t\}} \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] = \mathbb{E}[\mathbf{1}_{\{\tau \leq t\}} \mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] = 0.$$

From Lemma 3.5.2 the result follows by setting $Y = \mathbf{1}_{\{\tau > T\}}$, namely

$$\mathbf{1}_{\{\tau > t\}} \mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t^S] = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[\mathbf{1}_{\{\tau > T\}} \mathbf{1}_{\{\tau > t\}} | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)} = \mathbf{1}_{\{\tau > t\}} \frac{\mathbb{E}[\mathbf{1}_{\{\tau > T\}} | \mathcal{F}_t]}{\mathbb{P}(\tau > t | \mathcal{F}_t)}.$$

□

3.5.3 Proof of Lemma 3.3.6

The proof is along the same line as in JEANBLANC & VALCHEV [47, Lemma 2]. We have by the definition of the conditional probability

$$\mathbb{P}(M_t \leq u | X_t = v) = \frac{\mathbb{P}(M_t \leq u, X_t \in dv)}{\mathbb{P}(X_t \in dv)}.$$

A classical result from the reflection principle is (see HARRISON [39, Sec. 1.6])

$$\mathbb{P}(M_t \leq u, X_t > v) = \exp \left\{ \frac{2m}{\sigma^2} \ln \left(\frac{u}{x_0} \right) \right\} \Phi \left(\frac{-\ln(v/x_0) + 2\ln(u/x_0) + mt}{\sigma\sqrt{t}} \right).$$

Taking the derivative with respect to v leads to

$$\begin{aligned} \mathbb{P}(M_t \leq u, X_t \in dv) &= \varphi \left(\frac{-\ln(v/x_0) + 2\ln(u/x_0) + mt}{\sigma\sqrt{t}} \right) \frac{dv}{v\sigma\sqrt{t}} \times \\ &\quad \exp \left\{ \frac{2m}{\sigma^2} \ln \left(\frac{u}{x_0} \right) \right\}. \end{aligned}$$

For the distribution of X_t we have (see (3.13))

$$\mathbb{P}(X_t \leq v) = \Phi \left(\frac{\ln(v/x_0) - mt}{\sigma\sqrt{t}} \right)$$

and taking the derivative with respect to v yields

$$\mathbb{P}(X_t \in dv) = \varphi\left(\frac{\ln(v/x_0) - mt}{\sigma\sqrt{t}}\right) \frac{1}{v\sigma\sqrt{t}} dv.$$

The result follows from

$$\begin{aligned} & \frac{2m}{\sigma^2} \ln\left(\frac{u}{x_0}\right) - \frac{1}{2} \left(\frac{-\ln\left(\frac{v}{x_0}\right) + 2\ln\left(\frac{u}{x_0}\right) + mt}{\sigma\sqrt{t}} \right)^2 + \frac{1}{2} \left(\frac{\ln\left(\frac{v}{x_0}\right) - mt}{\sigma\sqrt{t}} \right)^2 \\ &= \frac{2m}{\sigma^2} \ln\left(\frac{u}{x_0}\right) - \frac{1}{2\sigma^2 t} \left[\left(\ln\left(\frac{v}{x_0}\right) \right)^2 - 2\ln\left(\frac{v}{x_0}\right) \left(2\ln\left(\frac{u}{x_0}\right) + mt \right) \right. \\ & \quad \left. + 4 \left(\ln\left(\frac{u}{x_0}\right) \right)^2 + 4\ln\left(\frac{u}{x_0}\right) mt + m^2 t^2 \right] \\ & \quad + \frac{1}{2\sigma^2 t} \left[\left(\ln\left(\frac{v}{x_0}\right) \right)^2 - 2\ln\left(\frac{v}{x_0}\right) mt + m^2 t^2 \right] \\ &= \frac{2}{\sigma^2 t} \left(\ln\left(\frac{v}{x_0}\right) \ln\left(\frac{u}{x_0}\right) - \left(\ln\left(\frac{u}{x_0}\right) \right)^2 \right) = \frac{-2}{\sigma^2 t} \ln\left(\frac{u}{x_0}\right) \left(\ln\left(\frac{u}{x_0}\right) - \ln\left(\frac{v}{x_0}\right) \right) \\ &= -\frac{2}{\sigma^2 t} \ln\left(\frac{u}{x_0}\right) \ln\left(\frac{u}{v}\right). \end{aligned}$$

3.5.4 Proof of Theorem 3.3.7

Observe that $\mathbb{F}^D \subset \mathbb{F}$. Then the proof of Eq. (3.10) is along the same line as the proof of Eq. (3.6) for the case of an S-investor. Indeed, from Proposition 3.5.1 we immediately obtain $\mathcal{G}_t^D \subset \mathcal{G}_t \subset \mathcal{K}_t$. Using the same arguments applied in the proof of Lemma 3.5.2 yields

$$\mathbb{1}_{\{\tau > t\}} \mathbb{E}[Y | \mathcal{G}_t^S] = \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{E}[Y \mathbb{1}_{\{\tau > t\}} | \mathcal{F}_t^D]}{\mathbb{P}(\tau > t | \mathcal{F}_t^D)},$$

where Y is an integrable random variable. Setting $Y = \mathbb{1}_{\{\tau > T\}}$ provides the result. For the sake of a simple notation we prove the remaining formulas of the theorem for the case $n = 2$, i.e., the threshold is L^1 in the interval $[t_0, t_1)$ and L^2 in the interval $[t_1, T)$. The proof for $n > 2$ is along the same line and skipped. In order to keep the notation simple we denote the times at which the D-investor observes the asset process by T_i , $i = 0, \dots, J-1$, where $T_0 := t_0$, $T_J := T$ and $T_j = t_1$ for some $j \in \{1, \dots, J-1\}$. The last term ensures that the D-investor obtains information about the asset process at the adjustment time t_1 of the default threshold. For any

$i \in \{1, \dots, J-1\}$ the process $(Y_u^i)_{u \geq 0}$ denotes a geometric Brownian motion defined by

$$Y_u^i = \exp\{mu + \sigma \widehat{B}_u^i\},$$

where $(\widehat{B}_u^i)_{u \geq 0}$ is a Brownian motion independent of \mathcal{F}_{T_i} and given by $\widehat{B}_u^i = B_{T_i+u} - B_{T_i}$. Note that Y_u^i inherits the independence of \mathcal{F}_{T_i} and it has the same law as X_u . Further we have the decomposition $X_s = X_{T_i} Y_{s-T_i}^i$ for $s > T_i$. We denote by $(\widehat{M}_u^i)_{u \geq 0}$ the running minimum process of $(Y_u^i)_{u \geq 0}$, i.e.,

$$\widehat{M}_u^i = \inf_{s < u} Y_s^i.$$

Explicit expressions for the joint density function $f_t^{\widehat{M}_t^i, Y_t^i}$ of Y_t^i and \widehat{M}_t^i and for the probability density function $f_t^{\widehat{M}_t^i}$ of \widehat{M}_t^i are provided in Lemma 3.3.1.

1. For the proof of (3.A) let $t \in [t_0, t_1)$ and $t \in [T_i, T_{i+1})$ with $i < j$, i.e., $T_{i+1} \leq T_j = t_1$. Then

$$\begin{aligned} \{\tau > T\} &= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, T]}\} \\ &= \{L^1 < M_{T_i}\} \cap \{L^1 < M_{[T_i, t_1]}\} \cap \{L^2 < M_{[t_1, T]}\} \\ &= \{L^1 < M_{T_i}\} \cap \{L^1 < M_{[T_i, T_j]}\} \cap \{L^2 < M_{[T_j, T]}\} \\ &= \{L^1 < M_{T_i}\} \cap \{L^1 < \widehat{M}_{T_j-T_i}^i X_{T_i}\} \cap \{L^2 < \widehat{M}_{T-T_j}^j Y_{T_j-T_i}^i X_{T_i}\}. \end{aligned}$$

Then we have

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t^D) &= \mathbb{P}(\tau > T | \mathcal{F}_{T_i}^D) \\ &= \mathbb{P}(L^1 < M_{T_i}, L^1 < \widehat{M}_{T_j-T_i}^i X_{T_i}, L^2 < \widehat{M}_{T-T_j}^j Y_{T_j-T_i}^i X_{T_i} | \mathcal{F}_{T_i}^D) \\ &= \int_0^1 \int_0^\infty \mathbb{P}(L^1 < M_{T_i}, L^1 < u X_{T_i}, L^2 < \widehat{M}_{T-T_j}^j v X_{T_i} | \mathcal{F}_{T_i}^D) \\ &\quad f_{T_j-T_i}^{\widehat{M}^i, Y^i}(u, v) dv du \\ &= \int_0^1 \int_0^\infty \int_0^1 \mathbb{P}(L^1 < M_{T_i}, L^1 < u X_{T_i}, L^2 < w v X_{T_i} | \mathcal{F}_{T_i}^D) \\ &\quad f_{T-T_j}^{\widehat{M}^j}(w) f_{T_j-T_i}^{\widehat{M}^i, Y^i}(u, v) dw dv du \\ &= \int_0^1 \int_0^\infty \int_0^1 \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_i}, \ell^1 < u X_{T_i}, \ell^2 < w v X_{T_i} | \mathcal{F}_{T_i}^D) \\ &\quad f_{L^1, L^2}(\ell^1, \ell^2) f_{T-T_j}^{\widehat{M}^j}(w) f_{T_j-T_i}^{\widehat{M}^i, Y^i}(u, v) d\ell^2 d\ell^1 dw dv du \\ &= \int_0^1 \int_0^\infty \int_0^1 \int_0^{u X_{T_i}} \int_0^{w v X_{T_i}} \mathbb{P}(\ell^1 < M_{T_i} | \mathcal{F}_{T_i}^D) \\ &\quad f_{L^1, L^2}(\ell^1, \ell^2) f_{T-T_j}^{\widehat{M}^j}(w) f_{T_j-T_i}^{\widehat{M}^i, Y^i}(u, v) d\ell^2 d\ell^1 dw dv du, \end{aligned}$$

where we have used in the second to last equation that (L^1, L^2) is independent of \mathcal{F}_T . Before we calculate the probability in the above integrand we introduce the following notation

$$P_i(\ell) = \mathbb{P}(\ell < M_{T_i} | X_{T_0}, \dots, X_{T_i}).$$

It holds

$$\begin{aligned} \{\ell < M_{T_i}\} &= \{\ell < \inf_{s < T_i} X_s\} = \{\ell < \inf_{s < T_{i-1}} X_s\} \cap \{\ell < \inf_{T_{i-1} \leq s < T_i} X_s\} \\ &= \{\ell < M_{T_{i-1}}\} \cap \{\ell < M_{[T_{i-1}, T_i]}\}. \end{aligned}$$

We have

$$\begin{aligned} P_i(\ell^1) &= \mathbb{P}(\ell^1 < M_{T_i} | X_{T_0}, \dots, X_{T_i}) = \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_i}\}} | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{1}_{\{\ell^1 < M_{[T_{i-1}, T_i]}\}} | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{1}_{\{\ell^1 < M_{[T_{i-1}, T_i]}\}} | X_s : s \leq T_{i-1}, X_{T_i}] | X_{T_0}, \dots, X_{T_i}], \end{aligned}$$

where the last equation follows from the tower property of the conditional expectation, since $\sigma(X_{T_0}, \dots, X_{T_i}) \subset \mathcal{F}_{T_{i-1}} \vee \sigma(X_{T_i})$. We obtain

$$\begin{aligned} P_i(\ell^1) &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{1}_{\{\ell^1 < M_{[T_{i-1}, T_i]}\}} | X_s : s \leq T_{i-1}, X_{T_i}] | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{[T_{i-1}, T_i]}\}} | X_s : s \leq T_{i-1}, X_{T_i}] | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{P}(\ell^1 < M_{[T_{i-1}, T_i]} | X_s : s \leq T_{i-1}, X_{T_i}) | X_{T_0}, \dots, X_{T_i}], \end{aligned}$$

where we have used that $M_{T_{i-1}}$ is $\mathcal{F}_{T_{i-1}}$ -measurable. The probability in the above equation is rewritten as

$$\begin{aligned} &\mathbb{P}(\ell^1 < M_{[T_{i-1}, T_i]} | X_s : s \leq T_{i-1}, X_{T_i}) \\ &= \mathbb{P}(\ell^1 < \inf_{T_{i-1} \leq s < T_i} X_s | X_s : s \leq T_{i-1}, X_{T_i}) \\ &= \mathbb{P}(\ell^1 < \inf_{T_{i-1} \leq s < T_i} X_{T_{i-1}} Y_{s-T_{i-1}}^{i-1} | X_s : s \leq T_{i-1}, X_{T_i}) \\ &= \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_s : s \leq T_{i-1}, X_{T_i}) \\ &= \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}), \end{aligned}$$

where the last equality holds since $\widehat{M}_{T_i - T_{i-1}}^{i-1}$ is independent of $\mathcal{F}_{T_{i-1}}$. Finally, we obtain the following recursion formula

$$\begin{aligned} P_i(\ell^1) &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{P}(\ell^1 < M_{[T_{i-1}, T_i]} | X_s : s \leq T_{i-1}, X_{T_i}) | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}) | X_{T_0}, \dots, X_{T_i}] \\ &= \mathbb{E}[\mathbb{1}_{\{\ell^1 < M_{T_{i-1}}\}} | X_{T_0}, \dots, X_{T_{i-1}}] \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}) \\ &= \mathbb{P}(\ell^1 < M_{T_{i-1}} | X_{T_0}, \dots, X_{T_{i-1}}) \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}) \\ &= P_{i-1}(\ell^1) \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}). \end{aligned}$$

Lemma 3.3.6 yields

$$\begin{aligned} & \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}) \\ &= 1 - \exp \left\{ \frac{-2}{\sigma^2(T_i - T_{i-1})} \ln \left(\frac{\ell^1}{X_{T_{i-1}}} \right) \ln \left(\frac{\ell^1}{X_{T_i}} \right) \right\} \end{aligned}$$

for $\ell^1 < \min(X_{T_{i-1}}, X_{T_i})$ and zero otherwise. Note that

$$K_i(\ell) = \begin{cases} 1 - \exp \left\{ \frac{-2}{\sigma^2(T_i - T_{i-1})} \ln \left(\frac{\ell}{X_{T_{i-1}}} \right) \ln \left(\frac{\ell}{X_{T_i}} \right) \right\}, & \text{for } \ell < \min(X_{T_{i-1}}, X_{T_i}) \\ 0, & \text{else.} \end{cases}$$

Then the probability $P_i(\ell^1)$ can be calculated recursively by

$$\begin{aligned} P_i(\ell^1) &= \mathbb{P}(\ell^1 < M_{T_i} | X_{T_0}, \dots, X_{T_i}) \\ &= P_{i-1}(\ell^1) \mathbb{P}(\ell^1 / X_{T_{i-1}} < \widehat{M}_{T_i - T_{i-1}}^{i-1} | X_{T_{i-1}}, X_{T_i}) \\ &= P_{i-1}(\ell^1) K_i(\ell^1) = P_{i-2}(\ell^1) K_{i-1}(\ell^1) K_i(\ell^1) = \prod_{j=1}^i K_j(\ell^1), \end{aligned}$$

since $P_0(\ell^1) = \mathbb{P}(\ell^1 < M_{T_0} | X_{T_0}) = \mathbb{P}(\ell^1 < X_{T_0} | X_{T_0}) = \mathbf{1}_{\{\ell^1 < X_{T_0}\}}$.

Eventually we obtain for $t \in [t_0, t_1)$ and $t \in [T_i, T_{i+1})$ that

$$\begin{aligned} \mathbb{P}(\tau > T | \mathcal{F}_t^D) &= \int_0^1 \int_0^\infty \int_0^1 \int_0^{uX_{T_i}} \int_0^{wvX_{T_i}} P_i(\ell^1) f_{L^1, L^2}(\ell^1, \ell^2) f_{T-T_j}^{\widehat{M}^j}(w) \\ & \quad f_{T_j - T_i}^{\widehat{M}^i, Y^i}(u, v) d\ell^2 d\ell^1 dw dv du. \end{aligned}$$

For the proof of (3.B) let $t \in [t_1, T)$ and $t \in [T_i, T_{i+1})$ with $i \geq j$, i.e., $t_1 = T_j \leq T_i$. Then

$$\begin{aligned} \{\tau > T\} &= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, T)}\} = \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T)}\} \\ &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i)}\} \cap \{L^2 < M_{[T_i, T)}\} \\ &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i)}\} \cap \{L^2 < \inf_{T_i \leq s < T} X_s\} \\ &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i)}\} \cap \{L^2 < \inf_{T_i \leq s < T} X_{T_i} Y_{s-T_i}^i\} \\ &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i)}\} \cap \{L^2 < X_{T_i} \widehat{M}_{T-T_i}^i\} \end{aligned}$$

and we obtain

$$\begin{aligned}
\mathbb{P}(\tau > T | \mathcal{F}_t^D) &= \mathbb{P}(\tau > T | \mathcal{F}_{T_i}^D) = \mathbb{P}(\tau > T | X_{T_0}, \dots, X_{T_i}) \\
&= \mathbb{P}(L^1 < M_{T_j}, L^2 < M_{[T_j, T_i]}, L^2 < \widehat{M}_{T-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) \\
&= \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_j}, \ell^2 < M_{[T_j, T_i]}, \ell^2 < \widehat{M}_{T-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) \\
&\quad f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
&= \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_j}, \ell^2 < M_{[T_j, T_i]} | X_{T_0}, \dots, X_{T_i}) \\
&\quad \mathbb{P}(\ell^2 / X_{T_i} < \widehat{M}_{T-T_i}^i) f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
&= \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_j}, \ell^2 < M_{[T_j, T_i]} | X_{T_0}, \dots, X_{T_i}) \Psi \left(T - T_i, \frac{\ell^2}{X_{T_i}} \right) \\
&\quad f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
&= \int_0^\infty \int_0^\infty \prod_{k=1}^j K_k(\ell^1) \prod_{k=j+1}^i K_k(\ell^2) \Psi \left(T - T_i, \frac{\ell^2}{X_{T_i}} \right) \\
&\quad f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
&= \int_0^\infty \int_0^\infty P_j(\ell^1) \prod_{k=j+1}^i K_k(\ell^2) \Psi \left(T - T_i, \frac{\ell^2}{X_{T_i}} \right) f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1,
\end{aligned}$$

where $\Psi(t, \cdot)$ is the complementary distribution function of M_t given in Lemma 3.3.2.

2. In the next step we calculate the survival probability until time t given \mathcal{F}_t^D .

For $t \in [t_0, t_1)$ and $t \in [T_i, T_{i+1})$ we have

$$\begin{aligned}
\mathbb{P}(\tau > t | \mathcal{F}_t^D) &= \mathbb{P}(L^1 < M_t | \mathcal{F}_t^D) = \mathbb{P}(L^1 < M_t | \mathcal{F}_{T_i}^D) = \mathbb{P}(L^1 < M_t | X_{T_0}, \dots, X_{T_i}) \\
&= \mathbb{P}(L^1 < M_{T_i}, L^1 < M_{[T_i, t]} | X_{T_0}, \dots, X_{T_i}) \\
&= \mathbb{P}(L^1 < M_{T_i}, L^1 < \widehat{M}_{t-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) \\
&= \int_0^\infty \mathbb{P}(\ell^1 < M_{T_i}, \ell^1 < \widehat{M}_{t-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) f_{L^1}(\ell^1) d\ell^1 \\
&= \int_0^\infty \mathbb{P}(\ell^1 < M_{T_i} | X_{T_0}, \dots, X_{T_i}) \mathbb{P}(\ell^1 / X_{T_i} < \widehat{M}_{t-T_i}^i) f_{L^1}(\ell^1) d\ell^1 \\
&= \int_0^\infty P_i(\ell^1) \Psi \left(t - T_i, \frac{\ell^1}{X_{T_i}} \right) f_{L^1}(\ell^1) d\ell^1.
\end{aligned}$$

The second to last equality follows by the independence of $\widehat{M}_{t-T_i}^i$ from T_i and the last equality holds since $\widehat{M}_{t-T_i}^i \stackrel{d}{=} M_{t-T_i}$.

For $t \in [t_1, T)$ and $t \in [T_i, T_{i+1})$ with $i \geq j$, i.e., $t_1 = T_j \leq T_i$, it holds

$$\begin{aligned}
 \{\tau > T\} &= \{L^1 < M_{t_1}\} \cap \{L^2 < M_{[t_1, t]}\} = \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, t]}\} \\
 &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i]}\} \cap \{L^2 < M_{[T_i, t]}\} \\
 &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i]}\} \cap \{L^2 < \inf_{T_i \leq s < t} X_s\} \\
 &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i]}\} \cap \{L^2 < \inf_{T_i \leq s < t} X_{T_i} Y_{s-T_i}^i\} \\
 &= \{L^1 < M_{T_j}\} \cap \{L^2 < M_{[T_j, T_i]}\} \cap \{L^2 < X_{T_i} \widehat{M}_{t-T_i}^i\}.
 \end{aligned}$$

Using the same arguments as above yields

$$\begin{aligned}
 \mathbb{P}(\tau > t | \mathcal{F}_t^D) &= \mathbb{P}(\tau > t | \mathcal{F}_{T_i}^D) = \mathbb{P}(\tau > t | X_{T_0}, \dots, X_{T_i}) \\
 &= \mathbb{P}(L^1 < M_{T_j}, L^2 < M_{[T_j, T_i]}, L^2 < \widehat{M}_{t-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) \\
 &= \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_j}, \ell^2 < M_{[T_j, T_i]}, \ell^2 < \widehat{M}_{t-T_i}^i X_{T_i} | X_{T_0}, \dots, X_{T_i}) \\
 &\quad f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
 &= \int_0^\infty \int_0^\infty \mathbb{P}(\ell^1 < M_{T_j}, \ell^2 < M_{[T_j, T_i]} | X_{T_0}, \dots, X_{T_i}) \mathbb{P}(\ell^2 / X_{T_i} < \widehat{M}_{t-T_i}^i) \\
 &\quad f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1 \\
 &= \int_0^\infty \int_0^\infty P_j(\ell^1) \prod_{k=j+1}^i K_k(\ell^2) \Psi\left(t - T_i, \frac{\ell^2}{X_{T_i}}\right) f_{L^1, L^2}(\ell^1, \ell^2) d\ell^2 d\ell^1.
 \end{aligned}$$

Chapter 4

Conclusions

This thesis proposes two approaches to improve financial risk management. Chapter 2 deals with market risk, i.e., the risk of a loss due to price movements in financial securities. In this chapter the classical portfolio problem of an investor, which consists of finding a strategy which optimally allocates the investor's wealth between different financial securities and consumption, is extended by incorporating dynamic risk measures to control of market risk. The resulting portfolio optimization problems under dynamic risk constraints are investigated in a continuous-time and discrete-time setting by using dynamic programming techniques. The corresponding dynamic programming equations are numerically solved by using a policy improvement algorithm in the continuous-time setting and a backward induction algorithm in the discrete-time setting. The loss in expected utility of intermediate consumption and terminal wealth caused by imposing a dynamic risk constraint is examined by numerical examples showing that the loss of portfolio performance is not too large while the risk is notably reduced. Further, time discretization effects are investigated by numerical illustrations indicating that the loss of portfolio performance resulting from imposing a risk constraint is typically bigger than the loss resulting from infrequent trading. Chapter 3 is concerned with credit risk, i.e., the risk of a loss arising from a counterparty failing to fulfill its contractually agreed financial obligations. In this chapter, the traditional structural model where default is triggered when the value of the firm's asset passes below a constant threshold is generalized by relaxing the assumption of a constant default threshold and modeling it as a random variable whose value is chosen by the management of the firm and dynamically adjusted to account for changes in the economy or the appointment of a new firm management. Investors on the market have no access to the value of the threshold and only anticipate the distribution of the threshold. Different information levels on the firm value are distinguished and explicit formulas for the conditional default probability given these information levels are derived. The conditional survival probabilities with respect to the accessible information and the associated credit yield spreads are compared by numerical illustrations which show that the information level has a considerable impact on the estimation of the

default probability and the associated credit yield spread. Investors who have perfect information on the value process of the firm learn about the default threshold, i.e., they learn that the default threshold must lie below the current running minimum of the firm value if default has not yet occurred. Thus, the larger the distance of the firm value to the running minimum the less likely is a default and the investors adjust their estimation of the default probability accordingly. The associated credit yield spreads are high, if the firm value is close to its running minimum and low otherwise. Especially, if the firm value is far above its running minimum just before maturity, the investors know that there will be no default in the next instance of time. Thus, investors do not demand a default risk premium and the credit spread is zero. This is different for investors who do not have full access to the value process of the firm. Investors who only obtain information about the firm value at certain dates cannot be certain about the firm value just before maturity and thus they demand a nonzero default risk premium. Furthermore, the credit spreads at maturity depend on their last observed firm value.

This thesis proposes two models to measure and manage market risk and credit risk. An important step for future research is to integrate market risk and credit risk management within a single model to facilitate financial risk management.

Bibliography

- [1] D. Akume. *Risk-Constrained Dynamic Portfolio Management*. Dissertation, University of Yaounde I., 2007.
- [2] D. Akume, B. Luderer, and R. Wunderlich. Dynamic shortfall constraints for optimal portfolios. *Surveys in Mathematics and its Applications*, **5**:135–149, 2010.
- [3] P. Artzner and F. Delbaen. Default risk insurance and incomplete markets. *Mathematical Finance*, 5:187 – 195, 12 2006.
- [4] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- [5] S. Basak and A. Shapiro. Value-at-risk-based risk management: Optimal policies and asset prices. *The Review of Financial Studies*, 14(2):371–405, 2001.
- [6] H. Bauer. *Measure and Integration Theory*. De Gruyter Studies in Mathematics. De Gruyter, 2001.
- [7] N. Bäuerle and U. Rieder. *Markov Decision Processes with Applications to Finance*. Universitext. Springer Berlin Heidelberg, 2011.
- [8] N. Bäuerle and U. Rieder. Control improvement for jump-diffusion processes with applications to finance. *Applied Mathematics & Optimization*, 65(1):1–14, Feb 2012.
- [9] N. Bäuerle, S.P. Urban, and L.A.M. Veraart. The relaxed investor with partial information. *SIAM Journal on Financial Mathematics*, 3(1):304–327, 2012.
- [10] J. Belinha. *Meshless Methods in Biomechanics: Bone Tissue Remodelling Analysis*. Lecture Notes in Computational Vision and Biomechanics. Springer International Publishing, 2014.
- [11] R.E. Bellman. *An Introduction to the Theory of Dynamic Programming*. Rand Corporation, 1953.

-
- [12] R.E. Bellman. *Dynamic Programming*. Dover Books on Computer Science. Dover Publications, 2013.
- [13] R.E. Bellman and S.E. Dreyfus. *Applied Dynamic Programming*. Princeton Legacy Library. Princeton University Press, 2015.
- [14] F. Black and J.C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, 31(2):351–367, 1976.
- [15] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [16] C. Blanchet-Scalliet, C. Hillairet, and Y. Jiao. Successive enlargement of filtrations and application to insider information. Preprint [hal-01259711](#), 2016.
- [17] C. Bluhm and L. Overbeck. *Structured Credit Portfolio Analysis, Baskets and CDOs*. Chapman and Hall/CRC Financial Mathematics Series. CRC Press, 2006.
- [18] T.J. Böhme and B. Frank. *Hybrid Systems, Optimal Control and Hybrid Vehicles: Theory, Methods and Applications*. Advances in Industrial Control. Springer International Publishing, 2017.
- [19] I.M. Bomze, G. Di Pillo, F. Schoen, V.F. Demyanov, R. Fletcher, and T. Terlaky. *Nonlinear Optimization: Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy, July 1-7, 2007*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2010.
- [20] R. Boudarel, J. Delmas, and P. Guichet. *Dynamic Programming and Its Application to Optimal Control*. Mathematics in Science and Engineering. Elsevier Science, 1971.
- [21] P. Boyle and W. Tian. Portfolio management with constraints. *Mathematical Finance*, 17:319–343, 02 2007.
- [22] U. Çetin, R.A. Jarrow, P.E. Protter, and Y. Yildirim. Modeling credit risk with partial information. *The Annals of Applied Probability*, 14(3):1167–1178, 2004.
- [23] W. Chen, Z.J. Fu, and C.S. Chen. *Recent Advances in Radial Basis Function Collocation Methods*. SpringerBriefs in Applied Sciences and Technology. Springer Berlin Heidelberg, 2013.
- [24] D. Cuoco, H. He, and S. Isaenko. Optimal dynamic trading strategies with risk limits. *Operations Research*, 56(2):358–368, 2008.
- [25] M.H.A. Davis. *Markov Models and Optimization*. Chapman and Hall/CRC Monographs on Statistics and Applied Probability. Routledge, 2018.

-
- [26] J.C. Duan. Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 4:155 – 167, 04 1994.
- [27] D. Duffie and D. Lando. Term structures of credit spreads with incomplete accounting information. *Econometrica*, **69**:633–664, 2001.
- [28] D. Duffie and K.J. Singleton. Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12(4):687–720, 1999.
- [29] D.J. Duffy. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. The Wiley Finance Series. Wiley, 2006.
- [30] J. Ericsson and J. Reneby. The valuation of corporate liabilities: Theory and tests. SSE/EFI Working Paper Series in Economics and Finance 445, 2001.
- [31] J. Ericsson and J. Reneby. Estimating structural bond pricing models. *The Journal of Business*, 78(2):707–735, 2005.
- [32] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control*. Stochastic Modelling and Applied Probability. Springer New York, 2012.
- [33] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Applications of mathematics. Springer, 2006.
- [34] R. Fletcher. *Practical Methods of Optimization*. Wiley, 2013.
- [35] A. Gabih, J. Sass, and R. Wunderlich. Utility maximization under bounded expected loss. *Stochastic Models*, 25(3):375–407, 2009.
- [36] K. Giesecke and L. R. Goldberg. The market price of credit risk : the impact of asymmetric information. Working paper, 2008.
- [37] L. Goldberg and K. Giesecke. Forecasting default in the face of uncertainty. *The Journal of Derivatives*, 12, 05 2004.
- [38] P. Grandits, F. Hubalek, W. Schachermayer, and M. Žigo. Optimal expected exponential utility of dividend payments in a brownian risk model. *Scandinavian Actuarial Journal*, 2007(2):73–107, 2007.
- [39] J.M. Harrison. *Brownian motion and stochastic flow systems*. John Wiley & Sons, 1985.
- [40] J.M. Harrison and A. Zeevi. Dynamic scheduling of a multiclass queue in the halfin-whitt heavy traffic regime. *Operations Research*, 52(2):243–257, 2004.
- [41] C. Hillairet and Y. Jiao. Credit risk with asymmetric information on the default threshold. *Stochastics*, **84(2-3)**:183–198, 2012.

- [42] K. Hinderer. *Foundations of non-stationary dynamic programming with discrete time parameter*. Lecture notes in operations research and mathematical systems. Springer-Verlag, 1970.
- [43] R.A. Howard. *Dynamic Programming and Markov Processes*. Published jointly by the Technology Press of the Massachusetts Institute of Technology and, 1960.
- [44] O.L.R. Jacobs. *An introduction to dynamic programming: the theory of multi-stage decision process*. Chapman and Hall Limited, 1967.
- [45] R.A. Jarrow and P.E. Protter. *Structural versus Reduced-Form Models: A New Information-Based Perspective*, chapter 6, pages 118–131. Wiley-Blackwell, 2015.
- [46] R.A. Jarrow and S. Turnbull. Credit risk: Drawing the analogy. *Risk Magazine*, 9(5), 1992.
- [47] M. Jeanblanc and S. Valchev. Partial information and hazard process. *International Journal of Theoretical and Applied Finance*, **8(6)**:807—838, 2005.
- [48] M. Jeanblanc, M. Yor, and M. Chesney. Mathematical methods for financial markets. *Finance*, 31(1):81–85, 2010.
- [49] P. Jorion. *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk*. McGraw-Hill Education, 2006.
- [50] R. Korn. *Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous Time*. World Scientific, 1997.
- [51] R. Korn and E. Korn. *Option Pricing and Portfolio Optimization: Modern Methods of Financial Mathematics*. Crm Proceedings & Lecture Notes. American Mathematical Society, 2001.
- [52] N.V. Krylov and A.B. Aries. *Controlled Diffusion Processes*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2008.
- [53] P. Lakner and W. Liang. Optimal investment in a defaultable bond. *Mathematics and Financial Economics*, **1(3-4)**:283–310, 2008.
- [54] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, **2**:99–120, 1998.
- [55] M. Leippold, F. Trojani, and P. Vanini. Equilibrium impact of value-at-risk regulation. *Journal of Economic Dynamics and Control*, 30(8):1277 – 1313, 2006.
- [56] J. Li and Y.T. Chen. *Computational Partial Differential Equations Using MATLAB*. Chapman & Hall/CRC Applied Mathematics & Nonlinear Science. CRC Press, 2008.

-
- [57] G.R. Liu and Y.T. Gu. *An Introduction to Meshfree Methods and Their Programming*. Springer Netherlands, 2005.
- [58] A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton series in finance. Princeton University Press, 2015.
- [59] R.C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–57, 1969.
- [60] R.C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373 – 413, 1971.
- [61] R.C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.
- [62] G. Monegato. A note on extended gaussian quadrature rule. *Mathematics of Computation*, 30(136):812–817, 1976.
- [63] R.B. Nelsen. *An Introduction to Copulas*. Springer Series in Statistics. Springer New York, 2007.
- [64] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006.
- [65] B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Universitext. Springer Berlin Heidelberg, 2007.
- [66] F. Pfützner. Dynamic risk constraints in utility maximization for multidimensional financial market models. Master’s thesis, Brandenburg University of Technology Cottbus-Senftenberg, Germany, 2017.
- [67] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.
- [68] T.A. Pirvu. Portfolio optimization under the value-at-risk constraint. *Quantitative Finance*, 7(2):125–136, 2007.
- [69] T.A. Pirvu, S. Moreno, and A. Réveillac. CRRA utility maximization under risk constraints. *Communications on Stochastic Analysis*, 7(2):203–225, 2013.
- [70] P.E. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2005.
- [71] W. Putschögl and J. Sass. Optimal consumption and investment under partial information. *Decisions in Economics and Finance*, 31(2):137–170, 2008.

- [72] I. Redeker. Portfolio optimization under dynamic risk constraints. Master's thesis, Brandenburg University of Technology Cottbus-Senftenberg, Germany, 2014.
- [73] I. Redeker and R. Wunderlich. Portfolio optimization under dynamic risk constraints: Continuous vs. discrete time trading. *Statistics & Risk Modeling*, 35(1-2):1–21, January 2018.
- [74] L.C.G. Rogers. The relaxed investor and parameter uncertainty. *Finance and Stochastics*, 5(2):131–154, Apr 2001.
- [75] A.A. Shardin and R. Wunderlich. Partially observable stochastic optimal control problems for an energy storage. *Stochastics*, 89(1):280–310, 2017.
- [76] Z. Shi and B.J.M. Werker. Short-horizon regulation for long-term investors. *Journal of Banking & Finance*, 36(12):3227 – 3238, 2012. Systemic risk, Basel III, global financial stability and regulation.
- [77] M. Sklar. *Fonctions de Répartition À N Dimensions Et Leurs Marges*. Université Paris 8, 1959.
- [78] J.W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Texts in Applied Mathematics. Springer New York, 2013.
- [79] N. Touzi. *Stochastic control problems, viscosity solutions and application to finance*. Publications of the Scuola Normale Superiore. Scuola Normale Superiore, 2002.
- [80] J.G. Wang and G.R. Liu. On the optimal shape parameters of radial basis functions used for 2-d meshless methods. *Computer Methods in Applied Mechanics and Engineering*, 191(23):2611 – 2630, 2002.
- [81] S.J. Wright. *Primal-Dual Interior-Point Methods*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 1997.
- [82] K.F.C. Yiu. Optimal portfolios under a value-at-risk constraint. *Journal of Economic Dynamics and Control*, 2004.
- [83] J. Yong and X.Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Stochastic Modelling and Applied Probability. Springer New York, 2012.

List of Figures

2.1	Effect of the VaR constraint on the value function and optimal portfolio-proportion strategy for the continuous-time problem.	29
2.2	Efficiency of a VaR-constrained investor relative to a Merton investor in continuous time for different relative bounds $\tilde{\varepsilon}(t, x) = \lambda x$	30
2.3	Effect of an absolute bound for the VaR constraint on the value function and optimal investment proportion strategy for the continuous-time problem.	31
2.4	Effect of different relative bounds, different terminal trading times and different risk constraints for the continuous-time problem.	32
2.5	Effect of the VaR constraint on the value function and optimal portfolio-proportion strategy for the discrete-time problem.	33
2.6	Efficiency for different bounds and for different risk measurement horizons.	35
2.7	Case of two stocks - Effect of the VaR constraint on the value function and optimal investment proportion strategy for different values of ρ	44
2.8	Case of two stocks - Effect of the VaR constraint on the value function, optimal investment proportion strategy, and efficiency for different values of ρ	45
2.9	Case of two stocks - Effect of the TCE constraint on the value function and optimal investment proportion strategy for different benchmarks.	47
2.10	Case of two stocks - Effect of the TCE constraint on the value function at $t = 0$ and optimal investment proportion strategy at $t = 0$ for different benchmarks.	48
2.11	Case of two stocks - Effect of the EL constraint on the value function and optimal investment proportion strategy for different relative bounds.	50
2.12	Case of two stocks - Effect of the EL constraint on the value function and optimal investment proportion strategy for different absolute bounds.	51
2.13	Case of two stocks - Effect of the EL constraint on the value function at $t = 0$ and optimal investment proportion strategy at $t = 0$ for different bounds.	52
3.1	Firm value, running minimum process and switching barrier.	74

3.2	Constant default threshold - no default case: Firm value and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.	86
3.3	Constant default threshold - default case: Firm value and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.	87
3.4	Constant default threshold - time horizon 5 years: Firm value and the associated conditional survival probabilities and credit yield spreads given the information of an S-investor and a D-investor.	88
3.5	Switching default thresholds - no default case: Firm value and the associated conditional survival probability and credit spread given the information of an S-investor for different values of θ	92
3.6	Switching default thresholds - no default case: Firm value and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.	93
3.7	Switching default thresholds - default case: Firm value and the associated conditional survival probability and credit spread given the information of an S-investor for different values of θ	94
3.8	Switching default thresholds - default case: Firm value and the associated conditional survival probabilities and credit spreads given the information of an S-investor and D-investors.	95

Notation

I. General notation

For any real numbers x and y : $x \wedge y = \min(x, y)$, $x^+ = \max(x, 0)$.

II. Sets

\mathbb{N} is the set of nonnegative integers, $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. \mathbb{R}^d denotes the d -dimensional Euclidian space. $\mathbb{R}_*^d = \{(x_1, \dots, x_d) | x_k \geq 0, k = 1, \dots, d\}$, $\mathbb{R}_+^d = \{(x_1, \dots, x_d) | x_k > 0, k = 1, \dots, d\}$. $\mathbb{R} = \mathbb{R}^1$. For all $\mathbf{x} = (x_1, \dots, x_d)$ in \mathbb{R}^d , we denote by $\|\cdot\|$ the Euclidian norm: $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^d (x_i)^2}$. $\mathbb{R}^{n \times d}$ is the set of real-valued $n \times d$ matrices ($\mathbb{R}^{n \times 1} = \mathbb{R}^n$). For all $\boldsymbol{\sigma} = (\sigma^{ij})_{i=1, \dots, d}^{j=1, \dots, n} \in \mathbb{R}^{d \times n}$, we denote by $\boldsymbol{\sigma}' = (\sigma^{ji})_{j=1, \dots, n}^{i=1, \dots, d} \in \mathbb{R}^{n \times d}$ the transpose matrix. The trace of a $n \times n$ matrix $\mathbf{a} = (a^{ij})_{j=1, \dots, n}^{i=1, \dots, n}$ is denoted by $tr(\mathbf{a}) = \sum_{i=1}^n a^{ii}$.

III. Functions and function spaces

The indicator function of a set A is denoted by

$$\mathbb{1}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

For any open set \mathcal{O} of \mathbb{R}^n we denote $\mathcal{C}^k(\mathcal{O})$ the space of all real-valued continuous functions f on \mathcal{O} with continuous derivatives up to order k . $C^0([0, T] \times \mathcal{O})$ is the space of all real-valued continuous functions f on $[0, T] \times \mathcal{O}$, with $0 < T < \infty$. $C^{1,2}([0, T] \times \mathcal{O})$ is the space of real-valued functions f on $[0, T] \times \mathcal{O}$ whose partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$, $1 \leq i, j \leq n$, exist and are continuous on $[0, T]$. The gradient vector and the Hessian matrix of a function $\mathbf{x} \rightarrow f(t, \mathbf{x}) \in \mathcal{C}^2(\mathcal{O})$ are denoted by $D_{\mathbf{x}}f$ and $D_{\mathbf{x}}^2f$, respectively.

IV. Integration and probability

$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. \mathbb{P} a.s. denotes “almost surely for the probability measure \mathbb{P} ” (the reference to \mathbb{P} is omitted when there is no ambiguity). $\sigma(\mathcal{G})$ is the smallest σ -field containing \mathcal{G} . For two sets A and B we denote by $\mathcal{G} = A \vee B$ the smallest σ -field containing A and B . $\mathbb{E}[X]$ is the expectation of the random variable X with respect to a probability \mathbb{P} initially fixed. $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation of X given \mathcal{G} .

V. Abbreviations

DPP: dynamic programming principle

HJB: Hamilton-Jacobi-Bellman

ODE: ordinary differential equation

PDE: partial differential equation

SDE: stochastic differential equation

w.l.o.g.: without loss of generality